Author:
Gibson, Mark

Title:
How Articulatory Suppression, Visual-Spatial Suppression and Private Speech Affect Planning in Autism Spectrum Disorder

How Articulatory Suppression, Visual-Spatial Suppression and Private Speech Affect Planning in Autism Spectrum Disorder

Mark Gibson

University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Master by Research in the Faculty of Life Science.

School of Psychological Science

October 2019

Word Count: 14,764

Abstract

Much scientific evidence suggests that typically developing (TD) children transition from supporting goal-directed behaviour with visual-spatial processing to supporting goal-directed behaviour with dialogic inner speech, at the age of 7. Alternatively, research suggests that children with Autism Spectrum Disorder (ASD) over the age of 7 do not use inner speech, however, it is unclear whether this transition is absent (atypical development) in individuals with ASD or merely delayed (delayed development). Two experiments utilised the Tower of London task (computerised), under the conditions of silence, private speech, articulatory suppression and visual-spatial suppression to test these hypotheses. Both accuracy (number of moves over the minimum) and time taken were recorded. Experiment 1 tested children with and without ASD between the ages of 7 and 12 and Experiment 2 tested adults with and without ASD. There were no significant differences in intelligence between the ASD and TD groups in either Experiment 1 or 2. Neither Experiment 1, Experiment 2 or a post-hoc analysis which compared the data of both experiments found any significant group differences or significant interactions between group and condition. No significant effects on accuracy or significant effects of articulatory suppression were found. Only significant negative effects of visual-spatial suppression (for adults and children) and private speech (for adults) on time were found. No significant correlations were found between age and performance (relative to baseline) under the dual tasks for TD children or children with ASD. A significant negative correlation was found between baseline performance and relative performance (for time) under articulatory suppression in TD adults only. This pattern of results does not fit with what would be predicted by either the atypical or delayed development models and contradicts the previous literature (calling its replicability into question).

Dedication and Acknowledgements

Firstly, I would like to thank my supervisors Professor Christopher Jarrold and Professor Melissa Allen for all their help, guidance and support. I am hugely grateful for their hard work this year. I would also like to thank the staff, students and parents at Venturers' Academy, Fosse Way School, St. Michael's on the Mount Primary School and St. George's Primary School for their cooperation with data collection. Finally, I would like to thank the members of the Bristol Cognitive Development Centre for allowing me to use their participant database and laboratories.

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

Table of contents

Tables

Experiment 1

Experiment 2

Post-hoc Analysis of Experiments 1 and 2

1. General Introduction

According to Vygotsky (1934/1987), inner speech is used by children above the age of 7 to structure goal-directed behaviour. His theory claims that inner speech develops from private speech (talking to yourself), which develops from early social communication. This would mean that inner speech is dialogic in nature (Fernyhough, 1996), i.e. conversational, complex, grammatical, and in the form of full sentences. A large body of research supports Vygotsky's (1934/1987) theory in typically developing (TD) children, showing that a reliance on inner speech during goal-directed behaviour appears after the age of 7, and suggests that before this children instead rely on visual-spatial processing (see Alderson-Day & Fernyhough, 2015 for a review). In contrast, research suggests individuals with ASD may not transition from relying on visual-spatial processing to relying on inner speech to support goal-directed behaviour at the age of 7 (see Williams, Bowler & Jarrold, 2012). However, it is unclear whether this transition occurs later in childhood (delayed development) or never occurs at all (atypical development), as findings have not been consistent (Holland & Low, 2010; Lidstone, Fernyhough, Meins & Whitehouse, 2009; Wallace, Silvers, Martin & Kenworthy, 2009; Whitehouse, Maybery & Durkin, 2006; William et al., 2012).

1.1 Private Speech Use in TD Children

Evidence for the development of private speech into inner speech comes from Winsler, Diaz, and Montero (1997), who showed that private speech becomes more task focused and more inaudible between the ages of 3 and 8 as it is internalised. Furthermore, McGonigle-Chalmers, Slater, and Smith (2014) found that, while private speech is more common and more task-irrelevant around others, it still occurs when children are alone. This indicates that, while private speech has a social aspect to it, it is also independent of this. The evidence from these studies challenges Piaget's (1959) alternative theory that private speech is a failed attempt to communicate and suggests that private speech starts as socially orientated but becomes increasingly less so as the child ages. This is what would be expected if both private speech and inner speech develop from early social communication as Vygotsky (1934/1987) stated.

Evidence has strongly supported Vygotsky's (1934/1987) claim that private speech is internalised around the age of 7 and is used to support goal-directed behaviour. For example, natural private speech improves planning ability (Winsler & Naglieri, 2003) and experimentally manipulated private speech improves task switching ability (Kray, Eber, & Karbach, 2008) performance in younger TD children but not older TD children (aged 5 to 7 and 7 to 9 respectively). This suggests that older children are using inner speech to structure cognition, and hence engaging in private speech makes no difference to their performance as they are already using linguistic strategies to support goal-directed behaviour. Conversely, younger children have not learnt to use inner speech and so private speech improves their performance as it is their only method of supporting cognition verbally.

1.2 Inner Speech Use in TD Children

More evidence to support Vygotsky's (1934/1987) theory comes from memory studies. The phonological similarity effect (PSE), where recall for lists of phonologically similar items is worse than recall for lists of phonologically dissimilar items (Gathercole, 1998), suggests memory maintenance is supported by inner speech. Initial experiments in this area, which calculated absolute PSE effects (i.e. recall in baseline minus recall in the phonological similarity condition), suggested that the PSE only appeared after the age of 7 (Halliday, Hitch, Lennon, & Pettipher, 1990). Alternatively, the visual similarity effect, where recall for lists of visually similar items is worse than recall for lists of visually dissimilar items, has been found to only occur before the age of 7 (Hitch, Woodin, & Baker, 1989). This suggests that memory maintenance switches from being visually supported to being verbally supported around this age. However, absolute PSE effects are limited by overall recall. Younger children have worse overall recall and so it is harder to find evidence of the PSE when calculating absolute PSE effects. Jarrold and Citroën (2013) ran four versions of the PSE task (employing either visual or verbal encoding with visual or verbal recall) on children aged 4 to 9 years old and calculated proportional PSE effects. The PSE was shown to be constant across all ages on three versions of the task. While the visual encoding with recall task showed a smaller proportional PSE in children under 7, Jarrold and Citroën (2013) explain this finding through floor effects (the data indicated this version of the task was the hardest with the lowest baseline recall). This suggests that inner speech may be used to support cognition from before the age of 7, challenging Vygotsky's (1934/1987) claim. However, floor effects cannot explain the disappearance of the VSE at the age of 7, so it is difficult to argue that some form of shift in strategy use for memory maintenance (from primarily visual to primarily verbal) may not occur around the age of 7. Evidence has shown that private speech seen between 4 and 7 years has been shown to correlate with the size of the proportional PSE for visually presented material in TD children (Al-Namlah, Fernyhough & Meins, 2006). This links private speech use to inner speech development in young children and, while suggesting inner speech use may be present before the age of seven, provides preliminary evidence for Vygotsky's (1934/1987) theory that inner speech supports cognition and goal-directed behaviour and develops from private speech.

Evidence of articulatory suppression (repeating word sequences during dependent variable tests) interfering with memory maintenance and goal-directed behaviour above the age of 7 only is further evidence of this shift to verbally supported cognition at this age. For example, articulatory suppression has been shown to impair recall, eliminate the PSE when information is both presented visually (pictures) and verbally (Cowan, Cartwright, Winterowd & Sherk, 1987; Ford & Silber, 1994; Hasselhorn & Grube, 2003; Hitch et al., 1989), and induce the VSE (Hitch et al., 1989) when information is presented visually (pictures), in those above the age of 7. The studies using verbal presentation did not look for the presence of the VSE and so it is unknown if it can be induced without visual stimuli. The effects of articulatory suppression on the PSE and the VSE may be because articulatory suppression blocks inner speech and the ability to encode items verbally. However, it is also possible that articulatory suppression made the task more difficult and so eliminated the PSE by introducing floor effects (Wang, Logie & Jarrold, 2016), but again floor effects cannot explain how the VSE was induced. Articulatory suppression has also been shown to impair the performance of children above 7 years of age on tests of task switching (Holland & Low, 2010; Whitehouse et al., 2006) and planning (Lidstone, Meins, & Fernyhough, 2010). This is further evidence of inner speech supporting goal-directed behaviour in children above the age of 7.

1.3 Visual-Spatial and Verbal Processing in ASD

The stages of development seen in TD individuals and the use of inner speech to support goal-directed behaviour may differ in individuals with Autism Spectrum Disorder (ASD). ASD is a disorder defined by social and communication deficits and the display of repetitive behaviours and restricted interests (American Psychiatric Association, 2013; World Health Organization, 1992). Evidence suggests individuals with ASD also display deficits in executive functions (cognitive mechanisms for goal-directed behaviour) such as working memory, planning, inhibition and cognitive flexibility (Baron-Cohen & Swettenham, 1997; Ozonoff, Pennington & Rogers, 1991; Pennington & Ozonoff, 1986), and that individuals with ASD self-report to think in pictures rather than words (Grandin, 1995; Hurlburt, Happé, & Frith, 1994). Therefore, it could be hypothesised that individuals with ASD might be worse at executive function tasks due to the fact their atypical early social communication means they have not developed inner speech adequately enough to utilise it as a tool to support goal-directed behaviour. This hypothesis is supported by evidence that children with Specific Language Impairment between the ages of 7 and 11 (who struggle to engage in early social communication due to atypical physical or neurological development which impairs speech and language use) show less internalised private speech and are unaffected by articulatory suppression during planning tasks (Lidstone, Meins & Fernyhough, 2012). Furthermore, unlike TD children (Campbell et al., 2013), performance mental age and not verbal mental age predicts flexibility (being asked to select items which differ on a trait and then being asked to select items which differ on a different trait) among children with ASD (Campbell et al., 2017). The same is true for the performance of adults with ASD on the Vygotsky block task (Constable, Ring, Gaigg & Bowler, 2017) where participants must learn which blocks of different shapes and colour fit into the four non-word categories through feedback and then suggest other ways the blocks can be categorised. This evidence implies that inner speech may not be used to support goal-directed behaviour and so may be absent in individuals with ASD as the quality and development of speech does not predict performance as it does in TD children. Evidence also suggests that TD children only show superior performance compared to children with ASD on executive function tasks which require complex rule use and non-verbal responses, as complex rules can be maintained with inner speech when no verbalisation is required (Russell, Jarrold, & Hood, 1999). Furthermore, individuals with ASD show superior performance on visual-spatial tasks compared to TD individuals (Frith & Happé, 1994; Soulières, Zeffiro, Girard & Mottron, 2011). For example, children with ASD show superior performance on embedded figures tasks compared to TD children (Frith & Happé, 1994) and adolescents/young adults with ASD show superior performance on mental rotation tasks compared to TD adolescents/young adults (Soulières, Zeffiro, Girard & Mottron, 2011). This superior visual-spatial processing seen in individuals with ASD could imply that ASD individuals use visual-spatial processing more (and thus are more practised) due to never transitioning to inner speech use. Alternatively, if the superior visual-spatial ability seen in individuals with ASD is innate it might lead to the transition never occurring. Further evidence implying this transition does not occur in individuals with ASD comes from neuroimaging studies. When completing verbal tasks children with high-functioning ASD (between the ages of 11 and 18) show more activation in visual areas and less activation in verbal areas than TD children (Sayhoun, Beliveau, Soulières, Schwartz & Moody, 2010). Furthermore, when completing Raven's progressive matrices adolescents and young adults with ASD show the same activation pattern (Soulières et al., 2009). Again, this suggests that speech is not being used to aid task completion in individuals with ASD, unlike in TD individuals. The fact some of these effects have replicated in adults with ASD suggests the presence of atypical development and not simply delayed development (i.e. this transition never takes place).

1.4 Inner Speech Use and Memory Maintenance in ASD

Evidence has not been clear on whether memory maintenance is supported by inner speech in children with ASD or not. Joseph, Steele, Meyer, and Tager-Flusberg (2005) found that while there was no difference in verbal memory span, children with ASD, unlike TD children, were no better at recognising new objects in sets when the objects where nameable than when they were not (i.e. real objects versus novel/made up objects). This suggests that the children with ASD were not encoding the objects verbally. Furthermore, Whitehouse et al. (2006) found a reduced picture superiority effect and a reduced word length effect for information in recall tasks for children with ASD. The picture superiority effect is where pictures of nameable objects are recalled better than the names themselves. This is supposedly because the former is encoded both visually and verbally while the latter is only encoded verbally (Paivio & Csapo, 1973), however, evidence has also been provided for other explanations such as pictures being more perceptually distinct (Nelson, Reed, & Walling, 1976). The word length effect is where objects with longer names are recalled less well than objects with shorter names, even when they are presented visually. Evidence that children with ASD show a reduced word length effect suggests that children with ASD do not verbally encode pictorial information, like TD children, during recall tasks and so are not using inner speech to support memory maintenance. Alternatively, Williams, Happé, and Jarrold (2008) found that the replacement of the VSE by the PSE did occur in children with ASD, just later than the age of 7. While the former two studies can fit with either the atypical or delayed development models, the results of Williams et al. (2008) support the latter model.

1.5 Inner Speech Use During Task-Switching and Planning in ASD

A number of studies have concluded that articulatory suppression may have no effect at all on task switching (Holland & Low, 2010; Whitehouse et al., 2006) or planning (Holland & Low, 2010; Wallace et al., 2009) in children with ASD, however, these studies have concluded this without sufficient statistical evidence. For example, Wallace et al. (2009) concluded that TD adolescents, and not ASD adolescents, were significantly affected by articulatory suppression on the original Tower of London task (based on post-hoc t-tests between conditions) but no significant interaction between condition and group was found in their data and so it cannot be concluded that their groups were reliably different. Williams et al. (2012) calculated Wallace et al.'s (2009) effect size as $d = 0.29$. This indicates that, not only was the interaction between group and condition not significant, the difference between performance under articulatory suppression between the two groups was relatively small. Furthermore, Whitehouse (2006) concluded that the task switching ability of children with ASD was not affected by articulatory suppression, however, a re-analysis of the data from by Lidstone et al. (2009) suggests that 60% of the ASD group were significantly affected by articulatory suppression and that those who were not affected were the youngest. Similarly, while Holland and Low (2010) concluded the same as Whitehouse (2006), age did moderately correlate (r = -.37) with the effect of articulatory suppression on performance (Williams et al., 2012); the older the child with ASD the more articulatory suppression affected performance. The reanalysed data of Whitehouse (2006) and Holland and Low (2010) supports the delayed development model (Lidstone et al., 2009; Williams et al., 2012).

Evidence of inner speech use being absent in individuals with ASD instead of delayed comes from Williams et al. (2012) who found that, while adults with ASD could use inner speech for recall tasks (as performance was affected by phonological similarity and articulatory suppression as in TD adults), they could not use inner speech for planning tasks (the five-disk Tower of London task). They concluded from this that perhaps individuals with ASD can use simpler monologic inner speech (not conversational and therefore not developed from early social communication) for memory maintenance but cannot use dialogic inner speech to support planning. Furthermore, it was found that the severity of communication deficits negatively correlated with the effect of articulatory suppression on planning in the ASD group, which supports Vygotsky's (1934/1987) theory. Finally, a post-hoc analysis showed that performance on the block subtask (a visual-spatial measure) of the Wechsler Adult Intelligence Scale positively correlated with performance in the baseline condition in participants with ASD only, suggesting a greater reliance on visual-spatial processing to support goal-directed behaviour. However, both latter findings are correlations and so no causality could be inferred as it could be the case that a third factor such as general intellectual impairments predicts both visual-spatial intelligence and planning ability.

1.6 The Current Study

The previous literature fits into two camps, one which suggests the use of inner speech to support goal-directed behaviour may be absent in individuals with ASD and one which suggests the transition from using visual-spatial processing to inner speech to support goal-directed behaviour is merely delayed in individuals with ASD. Surprisingly, very little robust evidence has been provided which attempts to assess the effects of visual-spatial suppression or private speech on planning in individuals with ASD. I therefore conducted a study which investigated the effects of articulatory suppression, visual-spatial suppression and private speech on planning ability in individuals with ASD and TD individuals. Experiment 1 sampled children while Experiment 2 sampled adults. The study was conducted on both adults and children to assess the development of planning strategies across the lifespan in individuals with ASD compared to TD individuals. Research into the effects of visual-spatial suppression or private speech on planning in individuals with ASD could help us further distinguish between the atypical and delayed development models. Understanding how individuals with ASD plan differently to TD individuals is important as it both increases our understanding of the disorder and can be used to create coping strategies for individuals with ASD who struggle in daily life (i.e. visualising or verbalising a problem).

2. Experiment 1

2.1 Introduction

The previous literature has largely supported Vygotsky's (1934/1987) theory that TD children switch from supporting goal-directed behaviour with visual-spatial processing to supporting goal-directed behaviour with inner speech at the age of 7 (see Alderson-Day & Fernyhough, 2015 for a review). Research has also indicated that this transition may not occur in children with ASD (Holland & Low, 2010; Wallace et al., 2009; Whitehouse et al., 2006, Williams et al., 2012), however, it is unclear whether this transition never occurs (atypical development) or is merely delayed (delayed development). While the effect of articulatory suppression on goal-directed behaviour in children with ASD has been heavily researched, few studies have researched the effects of visual-spatial suppression or private speech. A greater understanding of these effects could help us further distinguish between the atypical and delayed development models and improve our understanding of ASD.

2.1.1 Visual-Spatial Suppression in ASD. Little robust research has been conducted on the effects of visual-spatial suppression during goal-directed behaviour in ASD and the research that does exist is unable to help distinguish between either developmental models. Mitsuhashi, Shota, Hirata, Shogo, Okuzumi, and Hideyuki (2018) found that performance on a memory task (Luria hand task) was decreased equally by articulatory and visual-spatial suppression in children with ASD. This suggests that children with ASD both use inner speech to support planning and do not rely on visual-spatial processing more than inner speech. However, there was no control group to compare this to and so it could be the case that the tasks loaded general attention to different amounts and therefore no conclusion can be drawn from this evidence. Holland and Low (2010) found that visual-spatial suppression affected both the ASD group and control group equally for task switching and planning. This fits with neither the atypical nor delayed development model as both would predict children with ASD to rely more heavily on visual-spatial suppression (the difference being the latter model predicts this reliance will decrease with age while the former does not). However, Holland and Low (2010) used the Tower of Hanoi task which is much less versatile than the Tower of London task used by Williams et al. (2012) and Wallace et al. (2009), i.e. the difficulty of the puzzle can only be changed by adding or removing discs and there is only one puzzle/solution per difficulty/number of discs. Holland and Low (2010) repeated the same trials in each condition for every participant because of this, meaning group differences in the visual-spatial suppression condition may have been missed due to practice effects reducing power. However, a group difference was found for the effect of articulatory suppression on planning meaning that if there was a group difference in the effect of visual-spatial suppression on planning then this would have to be smaller than the group difference for the effect of articulatory suppression to not be detected. While a large amount of evidence implies individuals with ASD show visual-spatial advantages, it must also be remembered that visual-spatial deficits are also displayed (Schuh & Eigsti, 2012) and so it may not be the case that individuals with ASD rely on visual-spatial processing to support goal-directed behaviour more than TD individuals, even if they self-report to ‘think in pictures’ (Grandin, 1995; Hurlburt et al., 1994). It could be the case that visual-spatial advantages and the use of visual-spatial processing to support goal-directed behaviour within the ASD population vary with the level of general intellectual impairment or ASD symptomology. More research is needed to determine this.

2.1.2 Private Speech in ASD. The few experiments investigating the effects of private speech during goal-directed behaviour in ASD have not been able to provide consistent or robust findings. Hirata, Shogo, Okuzumi, Hideyuki, Kokubun, and Mitsuru (2016) found no difference in the performance of children with ASD on the Tower of Hanoi task when doing it alone in silence or having to instruct another to do it. However, there was no control group to compare to and arguably instructing another is a different and more challenging task than simply speaking your thoughts aloud while completing it alone. Russell-Smith, Comerford, Maybery, and Whitehouse (2014) found that articulatory suppression and encouraging private speech did not affect the performance on a card sorting task (task switching) of children with high-functioning ASD (aged 9 to 15), unlike TD controls. This suggests that children with ASD cannot use speech (either internally or externally) to support goal-directed behaviour like TD children do. This could be because they have not had sufficient early social communication to develop this strategy. This supports the atypical development model as if development was delayed in ASD, children with ASD would be expected to benefit from private speech like TD children under the age of 7 (who have not started to use inner speech) do. This challenges the findings of Winsler et al. (2007) who found that spontaneous private speech improved task switching in children with ASD (between the ages of 7 and 18) to a level that matched that of the TD group and that spontaneous private speech was equally common and equally developed (i.e. internalised and task focused) in both groups. This supports the delayed development model. However, Winsler et al. (2007) did not experimentally manipulate private speech but just compared those who spontaneously used it to those who did not. It could therefore be the case that the children with ASD who used private speech were less severely impaired both socially and cognitively and therefore performed better (i.e. the private speech did not improve performance but was a by-product of reduced symptom severity).

2.1.3 The Current Experiment. The effect of experimentally manipulated private speech during planning in children with ASD has yet to be properly investigated and the effect of visual-spatial suppression on planning in children with ASD has only been investigated once using a methodology that was subject to practice effects (Holland & Low, 2010). Experiment 1 attempted to investigate how articulatory suppression, visual-spatial suppression and (experimentally manipulated) private speech affect planning ability (on the Tower of London task) in children with ASD compared to TD children (above the age of 7) to address this gap in the previous literature and further distinguish between the delayed and atypical development models.

2.1.4 Hypotheses. If it is true that children with ASD never make the transition from visual-spatially supported goal-directed behaviour to inner speech supported goal-directed behaviour (the atypical development model) they will experience a greater negative impact on performance (time and accuracy – the larger the value the worse the performance) under visual-spatial suppression when compared to TD children, a reduced negative impact on performance under articulatory suppression when compared to TD children, and (like TD controls) show no benefit of private speech (for children above the age of 7). There will also be no significant correlations between age and relative performance (condition minus baseline condition) in any conditions for either group (i.e. their use of visual-spatial processing or inner speech does not change with age). This is because the strategies of neither group should change with age as the ASD group will make no transition to inner speech use and the TD group will have already transitioned completely to inner speech use. However, if the delayed development model is the more accurate model then there will be a positive effect for private speech in children with ASD and significant correlations between age and relative performance in the articulatory suppression, private speech (positive) and visual-spatial suppression (negative) conditions in the ASD group only (i.e. their reliance on inner speech increases, their benefit of private speech decreases and their reliance on visual-spatial processing decreases with age).

2.2 Methods

2.2.1 Participants. The final sample consisted of 14 TD children (Male = 9, Female = 5, Age Range = 7 years and 4 months to 11 years and 9 months, Mean Age = 9 years and 7 months) and 15 children with ASD (Male = 13, Female = 2, Age Range = 8 years and 0 months to 12 years and 11 months, Mean Age = 11 years and 4 months). An additional TD child was excluded during data collection due to being unable to understand English and therefore understand the task instructions, while 7 additional children with ASD were excluded due to being unable to complete the tasks. All children in the ASD group had previously had a formal diagnosis of ASD by a qualified specialist according to conventional criteria (American Psychiatric Association, 2013; World Health Organisation, 1992) and were sampled from special needs schools (except one who was sampled by invitation). All TD children were either sampled through primary schools or by invitation. Fully informed consent was obtained from parents/guardians prior to testing and ethical approval was obtained from the University of Bristol's School of Psychological Science Human Research Ethics Committee before data collection began (ethical approval code: 29111875781).

2.2.2 Materials and Stimuli. The children completed two tests of intelligence and one measure of planning ability (under four conditions), while the teachers or parents of the children completed a questionnaire which measured the children's ASD traits.

2.2.2.1 The British Picture Vocabulary Scale III (BPVS). was used to assess verbal intelligence (Dunn, Dunn & Styles, 2009). In this test children are told a word and must choose the corresponding picture out of a choice of four. The test takes approximately 10 minutes and produces a raw score which can be converted into a standardised score or a mental age.

2.2.2.2 Raven's Coloured Progressive Matrices (CPM). was used to calculate non-verbal intelligence (Raven, Raven & Court, 2003), and contains 3 sets of 12 questions (sets A, Ab and B) which become increasingly harder. In this task subjects must pick the element out of multiple options that completes the pattern (sometimes in the form of a matrix).

2.2.2.3 The Social Communication Questionnaire-Current (SCQ) was completed by the teachers or parents of the children, and measures the presence of ASD traits (Rutter, Bailey & Lord, 2003). The social communication questionnaire is made up of 40 items which are all true/false questions. The SCQ-Current asks about the child's behaviour in the last three months, unlike the SCQ-Lifetime which asks about the child's behaviour across their entire life (although the questions themselves are the same). The creators of the SCQ suggested a cut-off of 15 and found evidence to imply this correctly identified 96% of ASD individuals and 80% of TD individuals (Berument, Rutter, Lord, Pickles & Bailey, 1999). However, several studies have suggested a cut-off of 10-11 is needed for this level of accuracy (Barnard-Brak, Brewer, Chesnut, Richman & Schaeffer, 2016; Corsello et al., 2007; Eaves, Wingert, Ho, & Mickelson, 2006; Snow & Lecavalier, 2008).

2.2.2.4 The Tower of London Task measured planning ability and was based on the Five-Disc Tower of London task created by Ward and Allport (1997). In this task the participant must move five different coloured discs around on three equal sized pegs to create a new pattern (shown in the corner of the screen). This is different to the original Tower of London task where three coloured balls must be moved around on three unequal pegs (the first peg being able to hold three balls, the second two and the last one) to create a new pattern. First the basic rules and concepts of the Tower of London task were demonstrated using a 3D physical model. After this the participants started the task which was carried out on a touch screen tablet (Surface Pro 3) and was created in LiveCode (See Appendix A). The task was computerised as evidence suggests that computerised tasks produce more accurate measures of ability in individuals with ASD compared to equivalent non-computerised tasks, most likely due to the increased interest individuals with ASD display towards computers (Ozonoff, 1995). It was hoped that children with ASD would engage more with the computer task and be less likely to drop out due to frustration or boredom. The computerised task also reduces the likelihood of human error when recording the data. The task began with a set of instructions which explained that the participant had to help a cartoon character called "Felix the Fox" move coloured blocks around on three pegs to make the pattern he wants. The basic rules and concepts were also reiterated (i.e. only one block can be moved at a time and blocks cannot be moved from under other blocks or placed under blocks already on a peg). The instructions also stated they had to do this in as few moves as possible and as quickly as possible as well as indicating how many moves they thought each trial would take before they started moving blocks. The top line of a keyboard (which includes the numbers 0-9 and the backspace) was present at the bottom of the screen for them to input this estimate at the start of each trial. While this is not commonly an aspect of the Tower of London task this was added because Lidstone et al. (2010) only found an effect of articulatory suppression on the classic Five-Disc Tower of London task in children under the age of 10 when they had to announce the number of moves it would take to complete the trial before starting. This is most likely because this forced the children to consciously plan using working memory and not complete the task through trial and error or intuition. After reading the instructions the participant completed three practice trials. The children moved the blocks by tapping on them and then tapping on the unfilled spaces they wished to move the blocks to. Underneath a counter counted the number of moves they had taken. After completing the practice trials the participants completed 4 lists of 8 trials, each one under a different condition (See Appendix B). Each list contained the same number of trials of each difficulty level. Difficulty level was determined by the minimum number of moves to reach the goal state and the number of sub-goal moves (i.e. necessary moves which do not result in a block reaching its goal position). Ward and Allport (1997) found that the number of overall moves and the number of sub-goal moves significantly predicted time taken and errors made. The difficulty levels used were 2.0, 3.0, 3.1, 4.0, 4.1, 5.1, 5.2 and 6.1 where the first digit is the number of overall moves and the second is the number of sub-goal moves.

2.2.2.5 The Private Speech Score was a score between 0 and 5 rating the quality of the participants' private speech in the private speech condition. This was assessed by the experimenter during the condition. A score of 0 was given for no private speech, 1 for little private speech or non-task focused private speech, 2 for completely audible, loud and mostly task focused private speech, 3 for mostly audible, task focused private speech, 4 for mostly inaudible task focused private speech and 5 for completely inaudible private speech. This was based on the scoring scale of Winsler et al. (2007) who scored private speech as more developed the more task focused and internalised it was. This is in alignment with Vygotsky's (1934/1987) theory that private speech develops from early social communication and is internalised as the child ages (eventually developing into inner speech). This scoring system is supported by evidence that social private speech (louder, less task focused and displayed in the company of others) becomes less common as the child ages while asocial private speech (quieter and more task focused) becomes more common (McGonigle-Chalmers, Slater, and Smith, 2014; Winsler, Diaz, and Montero, 1997).

2.2.3 Design. The study used a mixed design. The between-subjects independent variable was whether the participant was TD or had ASD. The within-subjects independent variable was the condition of the Tower of London task (i.e. whether inner speech or visual-spatial processing was supressed or whether private speech was encouraged or discouraged), thus each participant received four conditions. The order of conditions and the list they were assigned to was counterbalanced between participants. The dependent variables were accuracy (how many more moves than the minimum number the participant took) and the amount of time (in seconds) to complete the whole trial (including the initial estimate). Non-verbal intelligence, verbal intelligence and ASD traits were measured to confirm the groups were matched in terms of intelligence and were different in terms of their ASD behaviours. The quality of private speech in the private speech condition was also measured to exclude participants who did not engage in sufficient private speech (i.e. no task focused private speech) from the analyses involving this condition and to look for group differences in private speech quality.

2.2.4 Procedure. Children either were tested in schools completing the tasks over three 25-minute sessions on different days (while the SCQ was completed by every child's teacher), or tested in their homes or at the University of Bristol during a single 1-hour session (while the SCQ was completed by their parent). For children being tested in schools (ASD = 14, TD = 4) the first session contained the measures used to determine their verbal and non-verbal intelligence, while the following two sessions contained two conditions of the Tower of London task each. For children being tested in a single session (ASD = 1, TD = 10) the Tower of London task was completed before the measures of verbal and non-verbal intelligence (participants were offered a break after the Tower of London task). The order of the Tower of London conditions was counterbalanced between participants. In the silence condition children were instructed to complete the task in silence. In the articulatory suppression condition the children were instructed to say 'Tuesday' in time with a metronome set to 65bpm. This is identical to the articulatory suppression condition used by Williams et al (2012). In the Visual-spatial suppression condition the children had to tap two raised plastic circles with one finger in time to a metronome set to 65bpm (in silence). The circle on the left was purple, 9.2cm in diameter and 0.8cm in height, while the circle on the right was teal, 5cm in diameter and 1.2cm in height. These circles were spaced 11.5cm apart from edge to edge and glued to a flat, green and A4 sized piece of plastic. A tapping task was used by Holland and Low (2010) to suppress visual-spatial processing, however, this task involved tapping four blocks in a sequence. The current method of visual-spatial suppression was chosen as it is more analogous to the articulatory suppression task which requires the participant to produce two syllables alternately. In the private speech condition, the child was told they must speak their thoughts aloud throughout the task. If the children stopped carrying out the dual tasks for several seconds the experimenter prompted them to continue. In all conditions but the visual-spatial suppression condition the children were instructed to only use one hand to complete the Tower of London task. This stopped children from performing better on silence, articulatory suppression and private speech conditions due to having both hands free and being able to complete the physical aspects of the tasks quicker. The metronome was played aloud in the silence and private speech conditions, but participants were instructed to ignore it.

2.2.5 Data Analysis. Preliminary independent samples t-tests were run to investigate differences in verbal intelligence, non-verbal intelligence, ASD traits and private speech quality between the two groups. The main analysis involved four mixed-design ANOVAs. The first two ANOVAs used group (two levels – whether the participant had received a formal diagnosis of ASD or not) as the between-subjects independent variable and condition (three levels – baseline, articulatory suppression and visual-spatial suppression) as the within-subjects independent variable with accuracy (moves over the minimum) as the dependent variable for one ANOVA and time (seconds) as the dependent variable for the other. The second two ANOVAs were identical except they used baseline and private speech as the conditions for the within-subjects independent variable (two levels) and excluded participants who had received a private speech score lower than 2. Separate ANOVAs to compare the private speech condition with the baseline condition excluding those who did not engage in sufficient private speech (i.e. no task focused private speech) were conducted to properly assess whether engaging in task focused private speech improved performance for either group (without reducing power for the comparison between the suppression conditions and the baseline condition). Pearson's correlations (Bonferroni corrected) were then run between age and relative performance (performance in condition minus performance in baseline) in the three non-baseline conditions for accuracy and time in each group (again excluding participants who had received a private speech score lower than 2 for correlations involving relative performance in the private speech condition).

2.3 Results

2.3.1 Preliminary t-tests. First preliminary independent samples t-tests were run to look for differences in verbal intelligence (BPVS raw scores), non-verbal intelligence (CPM scores), ASD traits (SCQ scores) and private speech scores between children with ASD and TD controls. There was no significant difference between BPVS score in the ASD (Mean = 124.27, SD = 23.06) and TD (Mean = 127.86, SD = 26.49) groups [$t(27) = 0.39$, $p = .700$, $d = 0.14$], between CPM score in the ASD (Mean = 31.47, SD = 5.14) and TD (Mean = 30.71, SD = 5.43) groups [$t(27) = -0.38$, $p = .704$, $d = 0.14$] or between private speech scores in the ASD (Mean = 2.47, SD = 1.51) and TD (Mean = 2.57, SD = 1.09) groups [$t(27) = 0.21$, $p = .833$, $d = 0.08$]. There was a significant difference between SCQ scores in the ASD (Mean = 14.57, SD = 4.97) and TD (Mean = 4.00, SD = 4.04) groups [$t(27) = -6.25$, $p < .001$, $d = 2.35$].

2.3.2 Primary Analysis (ANOVA). Next two mixed-design ANOVAs, using group (two levels) as the between-subjects independent variable, condition (three levels – baseline, articulatory suppression, visual-spatial suppression) as the within-subjects independent variable and accuracy (moves over the minimum) and time (seconds) as the dependent variables, were conducted. Mauchly's test of sphericity was not significant for either accuracy [$\chi2(2) = 0.73$, $p = .695$] or time [$\chi2(2) = 2.18$, $p = .336$] meaning assumptions of sphericity were not violated. There was no significant main effect of condition for accuracy [$F(2, 54) = 1.06$, $p = .351$] but there was a significant main effect of condition for time ($F(2, 54) = 13.79$, $p < .001$). There was no significant main effect of group for accuracy [$F(1, 27) = 0.05$, $p = .825$] or time [$F(1, 27) = 0.48$, $p = .491$] and no significant interaction between condition and group for accuracy [$F(2, 54) = 1.96$, $p = .151$] or time [$F(2, 54) = 0.31$, $p = .734$].

Post-hoc Bonferroni corrected t-tests conducted on time revealed that the visual-spatial suppression condition was the only condition significantly different to any others; participants were significantly slower in the visual-spatial suppression condition than in both the baseline condition [Mean Difference = 15.64, 95% CI = (8.72, 22.57), $p < .001$] and the articulatory suppression condition [Mean Difference = 11.03, 95% CI = (3.42, 18.63) $p = .003$]. Descriptive statistics can be seen in Table 1.

Table 1

Descriptive Statistics (Articulatory and Visual-Spatial Suppression). Means and Standard Deviations for the baseline, articulatory suppression and visual-spatial suppression conditions in each group in Experiment 1.

Measure	Condition	Group (N)	Mean	SD
Accuracy (Moves Over Minimum)	Baseline	ASD (15)	1.01	1.15
		TD (14)	0.71	0.60
	Articulatory Suppression	ASD (15)	0.57	0.67
		TD (14)	0.72	0.73
	Visual-spatial Suppression	ASD (15)	0.68	0.66
		TD (14)	0.98	0.76
Time (Seconds)	Baseline	ASD (15)	32.16	11.39
		TD (14)	34.50	14.51
	Articulatory Suppression	ASD (15)	34.71	10.83
		TD (14)	41.18	27.28
	Visual-spatial Suppression	ASD (15)	47.85	16.08
		TD (14)	50.10	17.95

Next two mixed-design ANOVAs, using group (two levels) as the between-subjects independent variable, condition (two levels – baseline and private speech) as the within-subjects independent variable and accuracy and time as the dependent variables, were conducted. For these ANOVAs participants who were given a private speech score below 2 were excluded from the analysis leaving 9 in the ASD group and 11 in the TD group. Levene's test indicted the assumption of equal variances was not violated for baseline accuracy [$F(1, 18) = 1.66$, $p = .213$], private speech accuracy [$F(1, 18) = 2.71$, $p = .117$], baseline time [$F(1, 18) = 0.05$, $p = .835$] or private speech time [$F(1, 18) = 1.12$, $p = .305$]. There was no significant main effect of condition for accuracy [$F(1, 18) = 0.07$, $p = .801$] or time [$F(1, 18) = 3.72$, $p = .070$], no significant main effect of group for accuracy [$F(1, 18) = 0.18$, $p = .673$] or time [$F(1, 18) = 0.06$, $p = .816$] and no significant interaction between condition and group for accuracy [$F(1, 18) = 0.80$, $p = .382$] or time [$F(1, 18) = 0.26$, $p = .616$]. Descriptive statistics can be seen in Table 2.

Table 2

Descriptive Statistics (Private Speech). Means and Standard Deviations for the baseline and private speech condition in each group after those who failed to engage in sufficient private speech were excluded in Experiment 1.

Measure	Condition	Group (N)	Mean	SD
Accuracy (Moves over Minimum)	Baseline	ASD (9)	0.83	0.89
		TD (11)	0.69	0.55
		Total (20)	0.76	0.70
	Private Speech	ASD (9)	0.63	0.68
		TD (11)	1.07	1.68
		Total (20)	0.87	1.31
Time (Seconds)	Baseline	ASD (9)	34.22	13.05
		TD (11)	31.73	13.34
		Total (20)	32.85	12.92
	Private Speech	ASD (9)	37.74	9.27
		TD (11)	33.78	14.21
		Total (20)	37.76	11.93

2.3.3 Correlations. Pearson's correlations were conducted to assess the relationships between age and relative performance (condition minus baseline) in the private speech, articulatory suppression and visual-spatial suppression conditions (for both accuracy and time) in each group, using a Bonferroni adjusted alpha level of .004 (.05/12). For the correlations between age and relative performance in the private speech condition participants who were given a private speech score below 2 were excluded. All correlations can be seen in Table 3. There were no significant correlations between age and relative performance in any conditions for either group.

Table 3.

Correlations. Pearson's correlations between age and relative performance in the non-baseline conditions for both accuracy and time in each group in Experiment 1.

Measure	Condition	Group	r	N	*P*
Accuracy (Moves Over Minimum)	Private Speech	ASD	-.56	9	.120
		TD	-.16	11	.637
	Articulatory Suppression	ASD	.19	15	.493
		TD	-.05	14	.859
	Visual-spatial Suppression	ASD	.01	15	.981
		TD	-.06	14	.836
Time (Seconds)	Private Speech	ASD	-.37	9	.333
		TD	-.51	11	.107
	Articulatory Suppression	ASD	.22	15	.421
		TD	-.23	14	.430
	Visual-spatial Suppression	ASD	.13	15	.640
		TD	.09	14	.749

2.4 Discussion

Both the atypical and delayed development models predict that children with ASD will use inner speech less, and visual spatial processing more that TD children when planning. The atypical development model also predicts neither group will benefit from private speech during planning and that these effects will not change with age. Alternatively, the delayed development model predicts that children with ASD will show a benefit of private speech during planning and that this benefit and the use of visual-spatial processing during planning will reduce with age while inner speech use during planning will increase with age.

Preliminary independent samples t-tests showed the groups were matched on verbal intelligence, non-verbal intelligence and private speech scores, while there was a significant difference in ASD traits between the two groups (children with ASD showed more ASD traits). The main analysis, four mixed-design ANOVAs, indicated there was no significant interaction between group and condition (for accuracy or time) and no significant main effect of group (on accuracy or time). However, the ANOVA comparing the baseline, articulatory suppression and visual-spatial suppression conditions on time found a significant main effect of condition with participants taking significantly longer in the visual-spatial suppression condition than the other two conditions. Finally, no significant correlations between age and relative performance were found.

2.4.1 Validity of Groups. There were no significant differences in intelligence (CPM and BPVS) between the two groups ($p < .700$) and the effect sizes were trivial ($d = 0.14$). There was also no significant difference in planning ability between the groups. The effect size for the differences in baseline performance between the two groups was small for accuracy ($d = 0.34$) and trivial for time ($d = 0.18$) and trivial for both once participants who did not engage in sufficient private speech were excluded ($d = 0.19$). This suggests the groups were equivalent in terms of cognitive ability.

A strong significant difference between the ASD and TD groups on SCQ scores was found ($p < .001$, $d = 2.35$) confirming that the lack of a significant group/condition interaction is not due to the ASD group containing many individuals with mild symptom severity or the TD group containing individuals with undiagnosed ASD. All of the TD group fell below the cut-off of 15 while 10 of the ASD group did. As 14 of the 15 ASD children were sampled from specialist schools it is unlikely that any have only mild symptom severity or have been misdiagnosed, suggesting this cut-off point given by the creators is inaccurate (Berument et al., 1999). When the alternative cut-off of 10 is used, all but two of the TD group fall below this cut-off while only two of the ASD group do. This supports several previous studies which have suggested the 10-11 as a more accurate cut-off (Barnard-Brak et al., 2016; Corsello et al., 2007; Eaves et al., 2006; Snow & Lecavalier, 2008). The small disparity seen here between formal diagnoses and SCQ scores could be due to the fact the SCQ-Current was used instead of the SCQ-Lifetime and it was filled out by teachers instead of primary care givers for the majority of the ASD sample. This is likely to make the SCQ scores less accurate.

Private speech scores were not significantly different between the two groups (p = .833, d = 0.08). However, 21% of the TD group were excluded from the ANOVAs comparing the private speech condition to baseline for not engaging in sufficient private speech, while 40% of the ASD group were excluded. This implies a difference in private speech quality may have existed which was not detected due to a lack of power. It is difficult to draw any conclusions about private speech/inner speech development in ASD from this large difference in the number of exclusions as it could be the case that differences in social anxiety between the two groups led to several members of the ASD group failing to speak their thoughts out loud. The comorbidity between ASD and anxiety has been well documented (Spain, Sin, Linder, McMahon & Happé, 2018).

2.4.2 Previous Literature. These results fail to fit with either the delayed or atypical development models. While the lack of significant correlations matches with what the atypical development model would predict, the lack of reliable group differences in the ANOVAs fit with neither model and provides no evidence that children with ASD are different to TD children when it comes to cognitive planning strategies. This challenges most of the previous literature which found that articulatory suppression impairs performance for TD children but not for children with ASD during goal-directed behaviour, specifically task switching (Holland & Low, 2010; Russell-Smith et al., 2014; Whitehouse et al., 2006) and planning (Holland & Low, 2010; Wallace et al., 2009). It could be possible that the lack of results is due to the sample size used in the experiment, which is small when compared to Wallace et al. (2009) and Whitehouse et al. (2006) who sampled between 20 and 30 participants per group. However, Holland and Low (2010) found the same pattern of results with a smaller sample size (13 in each group) and Russell-Smith et al (2014) also found the same pattern of results with sample sizes not much larger than that of the current study (15 to 17). Furthermore, the effect size for the group/condition interaction on planning found by Holland and Low (2010) was large (η_p^2=.46) indicating the interaction should be detected with a small sample size.

2.4.3 Conclusion. No significant effect of articulatory suppression and no significant interaction between group and condition was found. These results provide no evidence that suppressing inner speech has an effect on planning ability in children and that children with ASD are different in their use of inner speech or visual-spatial processing during planning compared to TD children. This pattern of results fails to fit with either the delayed or atypical development models and contradicts much of the previous literature. Research with adult participants could possibly do more to distinguish between the delayed and atypical development models.

3. Experiment 2

3.1 Introduction

Vygotsky (1934/1987) theorised that inner speech is used to support goal-directed behaviour in children above the age of 7 and evidence for this is mostly consistent (see Alderson-Day & Fernyhough, 2015 for a review). While there is a large body of evidence to indicate the use of inner speech is either absent or delayed in children with ASD (Holland & Low, 2010; Wallace et al., 2009; Whitehouse et al., 2006) very little research has been conducted using adults with ASD.

3.1.1 The Current Experiment. Currently, Williams et al. (2012) is the only previous study to investigate the effects of articulatory suppression on adults with and without ASD. It was found that the differences observed between TD children and children with ASD were maintained into adulthood (atypical development) but more research is needed to confirm this as a number of studies on children with ASD (Holland & Low, 2010; Lidstone et al., 2009; Whitehouse et al., 2006) have indicated these differences reduce with age (delayed development). Furthermore, no study has of yet investigated the effects of private speech and visual-spatial suppression on goal-directed behaviour in adults with and without ASD. As the different patterns of behaviour predicted by the delayed and atypical development models are more distinct in adulthood, research using adult participants is vital for distinguishing between the two models further and increasing our understanding of ASD. To address this gap in the literature Experiment 2 replicated Experiment 1 using adults.

3.1.2 Hypotheses. If the atypical development model is accurate then adults with ASD will show a greater impairment in performance under visual-spatial suppression when compared to TD adults, a reduced impairment in performance under articulatory suppression when compared to TD adults, and (like TD adults) show no benefit of private speech. There would also be a significant negative correlation between baseline performance (planning ability) and relative performance in the visual-spatial suppression condition for the ASD group, while there will only be a significant negative correlation between baseline performance and relative performance in the articulatory suppression condition for the TD group (i.e. adults with ASD will perform better when they rely more heavily on visual-spatial processing and TD adults will perform better when they rely more heavily on inner speech). Alternatively, if the delayed development model is more accurate then there should be no interaction between condition and group and there should only be significant negative correlations between baseline performance and relative performance in the articulatory suppression condition for both groups.

3.2 Methods

3.2.1 Participants. During data collection 2 participants were excluded for being given diagnoses of ASD as children which were overturned by clinicians during adolescence. The final samples consisted of 11 adults who had previously been diagnosed with ASD by a qualified professional (Male = 4, Female = 6, Non-Binary = 1) and 22 TD adults (Male = 12, Female = 10). All participants were sampled through the university and participated in exchange for course credit (TD only) or were sampled through advertising/invitation and participated in exchange for a £10 reimbursement (ASD only). Fully informed consent was obtained prior to testing and ethical approval was obtained from the University of Bristol's School of Psychological Science Human Research Ethics Committee before data collection began (ethical approval code: 29111875781).

3.2.2 Materials and Stimuli. Participants completed a measure of non-verbal intelligence, one measure of ASD traits and a version of the measure of planning ability used in Experiment 1. No measure of verbal intelligence was completed in Experiment 2. This was only measured in Experiment 1 as the participants were only just past the age of transitioning to inner speech use according to Vygotsky's (1934/1987) theory and any differences in verbal intelligence between the groups could have confounded the results. In adults without severe intellectual or verbal impairments it can be assumed that differences in inner speech use would not be caused by differences in verbal intelligence.

3.2.2.1 Raven's Advanced Progressive Matrices (APM) was used to asses non-verbal intelligence (Raven, Raven & Court, 2003). Unlike the CPM, it is targeted at more able adolescents and adults, contains 2 sets (the first being made up of 12 questions and the second being made of up 36) and is printed in black and white. Participants were given 40 minutes to complete as many answers from both sets as they could. The raw scores from each set were then combined to create a total score.

3.2.2.2 The Autism Quotient (AQ) is made up of 5 sets of 10 questions each measuring a different trait of ASD (Baron-Cohen, Wheelwright, Skinner, Martin & Clubley, 2001). These are social skills, attention switching, attention to detail, communication and imagination. The participants had to rate how much they agree with each statement (strongly agree, agree, disagree or strongly disagree). Participants receive one point for agreeing with a statement representing an ASD trait or disagreeing with a statement representing a lack of an ASD trait, otherwise they receive no point. Baron-Cohen et al. (2001) state that a score of 32 or above indicates "clinically significant levels of autistic traits". Woodbury-Smith, Robinson, Wheelwright and Baron-Cohen, (2005) have shown the discriminative validity of the AQ is fairly strong (classifying 77% of ASD individuals and 74% of TD individuals correctly) but suggested a cut-off of 26 for those referred for ASD screening by a clinician (classifying 95% of ASD individuals and 52% of TD individuals correctly).

3.2.2.3 *The Tower of London Task* was identical to that of Experiment 1 apart from the fact that participants completed four lists of 10 trials, each made up of the difficulty levels 3.0, 3.1, 4.0, 4.1, 5.1, 5.2, 6.2, 6.3, 7.2 and 7.3.

3.2.2.4 The Private Speech Score was calculated in the same way as in Experiment 1.

3.2.3 Design. The design was identical to that of Experiment 1, except verbal intelligence was not measured.

3.2.4 Procedure. Participants were tested individually in a computer lab during a one hour and forty-five-minute session. Participants completed the Tower of London task before the APM and AQ. Participants were offered a break after completing the Tower of London task. The conditions and procedure of the Tower of London task were identical to Experiment 1.

3.2.5 Data Analysis. The data analysis was identical to that of Experiment 1 except no t-test was run to look for differences in verbal intelligence between the two groups as there was no measure of verbal intelligence used in this experiment. Also, only two ANOVAs were conducted with condition (all four conditions) as the within-subjects independent variable, with accuracy and time as the dependent variables respectively. This is because no participant in either group received a private speech score lower than 2. Finally, Pearson's correlations were conducted between baseline performance (instead of age) and relative performance in the three non-baseline conditions.

3.3 Results

3.3.1 Preliminary t-tests. First preliminary t-tests were run to look for differences in non-verbal intelligence (APM scores), ASD traits (AQ scores) and private speech scores between adults with ASD and TD controls. There was no significant difference between APM scores in the ASD (Mean = 31.27, SD = 8.03) and TD (Mean = 33.45, SD =

1.67) group [t(31) = 0.75, p = .460, d = 0.45]. There was a significant difference between AQ scores in the ASD (Mean = 36.73, SD = 8.08) and TD (Mean = 12.59, SD = 5.28) group [t(31) = -10.35, p < .001, d = 3.61] and between private speech scores in the ASD (Mean = 2.55, SD = 0.69) and TD (Mean = 3.32, SD = 0.84) group [t(31) = 2.64, p = .013, d = 1.01].

3.3.2 Primary Analysis (ANOVA). Next two mixed-design ANOVAs, using group (two levels) as the between-subjects independent variable, condition (four levels) as the within-subjects independent variable and accuracy and time as the dependent variables, were conducted. Mauchly's test of sphericity was not significant for either accuracy [$\chi 2(5) = 1.27$, p = .938] or time [$\chi 2(5) = 6.33$, p = .276] meaning assumptions of sphericity were not violated. There was no significant main effect of condition for accuracy [F(3, 93) = 2.12, p = .103] but there was a significant main effect of condition for time [F(3, 93) = 11.20, p < .001]. There was no significant main effect of group for accuracy [F(1, 31) = 0.20, p = .655] or time [F(1, 31) = 0.06, p = .802] and there was no significant interaction between condition and group for accuracy [F(3, 93) = 1.06, p = .372] or time [F(3, 93) = 1.22, p = .308].

Post-hoc Bonferroni corrected t-tests conducted on time revealed that only the private speech condition [Mean Difference = 5.38, 95% CI = (2.00, 8.75), p = .001] and the visual-spatial suppression condition [Mean Difference = 7.89, 95% CI = (4.39, 11.39), p < .001] were significantly different to the baseline condition. Participants took significantly longer in both conditions compared to baseline. No other conditions were significantly different to each other. Descriptive statistics can be seen in Table 4.

Table 4

Descriptive statistics. Means and Standard Deviations for each condition and group in Experiment 2.

Measure	Condition	Group (N)	Mean	SD
Accuracy (Moves Over Minimum)	Baseline	ASD (11)	0.57	0.52
		TD (22)	0.81	1.15
	Private Speech	ASD (11)	0.98	0.63
		TD (22)	0.99	1.25
	Articulatory Suppression	ASD (11)	0.62	0.55
		TD (22)	0.97	1.11
	Visual-spatial Suppression	ASD (11)	0.87	0.75
		TD (22)	0.88	1.02
Time (Seconds)	Baseline	ASD (11)	26.05	7.87
		TD (22)	28.91	7.74
	Private Speech	ASD (11)	31.34	9.25
		TD (22)	33.67	10.01
	Articulatory Suppression	ASD (11)	31.38	11.69
		TD (22)	28.92	5.47
	Visual-spatial Suppression	ASD (11)	34.71	7.21
		TD (22)	35.33	9.05

3.3.3 Correlations. Pearson's correlations were conducted to assess the relationships between baseline performance and relative performance (condition minus baseline) in the private speech, articulatory suppression and visual-spatial suppression conditions (for both accuracy and time) in each group, using a Bonferroni adjusted alpha level of .004 (.05/12). All correlations can be seen in Table 5. There was a significant, strong, negative correlation between baseline performance and relative performance in the articulatory suppression condition (r = -.73, n = 22, $p < .001$) for time in the TD group. No other correlations were significant.

Table 5.

Correlations. Pearson's correlations between baseline performance and relative performance in the other conditions for both accuracy and time in each group in Experiment 2.

Measure	Condition	Group	r	N	*p*
Accuracy (Moves Over Minimum)	Private Speech	ASD	-.32	11	.333
		TD	-.14	22	.545
	Articulatory Suppression	ASD	-.60	11	.049
		TD	-.33	22	.135
	Visual-spatial Suppression	ASD	-.67	11	.025
		TD	-.47	22	.028
Time (Seconds)	Private Speech	ASD	-.36	11	.272
		TD	.15	22	.495
	Articulatory Suppression	ASD	-.08	11	.816
		TD	-.73	22	< .001
	Visual-spatial Suppression	ASD	-.58	11	.061
		TD	-.16	22	.465

3.4 Discussion

The atypical development model predicts that adults with ASD will use inner speech less, and visual spatial processing more that TD adults when planning. It also predicts that inner speech use will correlate with planning ability in TD adults and the use of visual-spatial processing will correlate with planning ability in adults with ASD. Alternatively, the delayed development models predict that adults with ASD will (like TD controls) solely use inner speech over visual-spatial processing during planning and that inner speech use will correlate with planning ability in both groups. Both models predict that neither group will show any benefit of private speech during planning.

Preliminary independent samples t-tests showed the groups were matched on non-verbal intelligence while there was a significant difference in private speech scores and ASD traits between the two groups (adults with ASD showed lower quality private speech and more ASD traits). The main analysis, two mixed-design ANOVAs, indicated there was no significant interaction between group and condition (for accuracy or time) and no significant effect of group (on accuracy or time) but there was a meaningful effect of condition (for time only) with participants taking significantly longer than the baseline condition in the private speech and visual-spatial suppression conditions. The only significant correlation found was a strong negative correlation between baseline performance and relative performance in the articulatory suppression condition for time in the TD group. This means as TD adults were quicker on the baseline condition, they were slower on the articulatory suppression condition compared to baseline (i.e. articulatory suppression time minus baseline time became more positive/less negative).

3.4.1 Validity of Groups. No significant difference in intelligence (APM) was found between the two groups (p = .460). However, the effect size (d = 0.45) was moderate indicating the ASD group had a slightly lower average intelligence. There were also small to moderate effect sizes for baseline planning ability (d = 0.29 for accuracy and d = 0.37 for time), however, these were in the opposite direction to intelligence (the ASD group performed slightly better). This indicates there may have been small differences in intelligence and planning ability between the two groups, however, these are unlikely to have affected the results (especially as the differences are not consistent in their direction).

A strong significant difference between the ASD and TD groups on AQ scores was found (p < .001, d = 3.61) confirming that the lack of a significant group/condition interaction is not due to the ASD group containing many individuals with mild symptom severity or the TD group containing individuals with undiagnosed ASD. All of the TD group fell below Baron-Cohen et al.'s (2001) cut-off point of 32, while only three of the ASD group did. All but one of the TD group fell below Woodbury-Smith et al.'s (2005) cut-off point of 26 while only one of the ASD group fell below this score.

A significant difference in private speech quality was also found (p = .013, d = 1.01). However, it is also difficult to draw conclusions from this group difference as the level of private speech engaged in by the participants is not a true representation of their private speech/inner speech development but is heavily dependent on each participant's interpretation of the instruction to speak their thoughts out loud. Perhaps adults with ASD interpreted the instruction more literally than TD adults, meaning their private speech was louder and less 'private'. Research has shown individuals with ASD often interpret language too literally (Kalandadze, Norbury, Nærland & Næss, 2018).

3.4.2 Previous Literature. This pattern of results fails to fit with either the delayed or atypical development models. While the delayed development model would predict no group differences between adults with ASD and TD adults and no evidence of group differences were found in the ANOVA, no significant effect of articulatory suppression on planning performance was found for either group which is contrary to what the model predicts. This is also contrary to the finding of Williams et al. (2012) who found adults with ASD were less impaired by articulatory suppression compared to TD controls during planning. It is possible that group differences were not found due to the small ASD sample of 11. This is slightly smaller than the samples used by Williams et al., (2012) which were 15 and 17. However, the effect size for the group/condition interaction found by Williams et al (2012) was moderate (η_p^2=.16), indicating that (if it is a true effect) it is unlikely that it wasn't found due to small sample sizes.

Private speech was found to slow participants. This contradicts previous literature which has suggested TD children above the age of 7 are unaffected by private speech (Kray et al., 2008; Winsler & Naglieri, 2003; Winsler et al., 2007). The current finding suggests that by adulthood individuals with and without ASD can think through problems in their heads (either verbally or visually) quicker than they can verbalise those thoughts out loud. While private speech was found to be less developed in the ASD group it is unlikely that this confounded the results as all participants received a score of 2 or above. This means that all participants were engaging in constant task focused private speech and any variation in scores was due to how internalised (i.e. loud and audible) it was.

A negative correlation was found between baseline performance and relative performance in the articulatory suppression condition for time in the TD group. This also contradicts the delayed development model which predicts this correlation would be found in both groups. This correlation instead fits with the atypical development model. However, no reliable correlation between baseline performance and relative performance in the visual-spatial suppression condition for the ASD group (as the atypical development model also predicts) was found.

3.4.3 Inner Speech Use in TD Adults. The current results contradict previous studies which found an effect of articulatory suppression during planning on TD adults. For example, Williams et al. (2012) found TD adults were impaired by articulatory suppression on memory and planning tasks. Furthermore, Baddeley, Chincotta and Adlam (2001) found over seven experiments that TD adults were impaired by articulatory suppression during task switching and Gilhooly (2005) found that articulatory suppression and not visual-spatial suppression impaired performance on a verbal reasoning tasks for TD adults. However, several studies investigating the effects of articulatory suppression on TD adults have indicated that they do not rely on inner speech to support goal-directed behaviour. Miyake, Emerson, Padilla, and

Ahn (2004) found that articulatory suppression only impaired task switching in TD adults when the switch cues (cues indicating a change in task) were the first letter of a word instead of the word itself (i.e. S or C instead of SHAPE or COLOUR) which forced the participants to use inner speech to complete the cue. Furthermore, Wynn, Gilhooly, Della Sala, and Logie (1999) found no evidence that articulatory suppression increased the number of moves made by TD adults during the Tower of London task and Gilhooly, Wynn, Phillips, Logie, and Della Sala (2002) found that planning ability in TD adults related to visual-spatial abilities and not verbal abilities. Finally, Law, Trawley, Brown, Stephens and Logie (2013) found, contrary to Gilhooly (2005), no effect of articulatory suppression on a virtual reality planning task but an effect of concurrent auditory localisation (visual-spatial suppression). It was concluded from this that TD adults rely more on visual-spatial processing than inner speech to support cognition, however, it could be hypothesised that auditory localisation requires what could be referred to as 'internal silence' and thus suppresses both inner speech and visual-spatial processing. Hence it may not be that TD adults are more reliant on visual-spatial processing for supporting goal-directed behaviour, just that TD adults rely on inner speech by default but can also use visual-spatial processing when inner speech is suppressed. An alternative explanation which could also explain the disparity between Gilhooly (2005) and Law et al. (2013) is that the task in the former study was completely verbal while the task in the latter was heavily visual-spatial; perhaps TD adults display a division of labour between verbal and visual-spatial processing when supporting goal-directed behaviour. Evidence of TD adults not using inner speech during planning and evidence of their planning strategies being more task dependent than in children could possibly explain why no significant group/condition interaction and no significant effect of articulatory suppression on the TD group was found in this experiment. It might be argued that the Tower of London task is a very visual-spatial task, however, Williams et al. (2012) also used the Tower of London task and found an effect of articulatory suppression in TD adults meaning this hypothesis cannot explain the disparity between this experiment and Williams et al. (2012).

3.4.4 Conclusion. No significant effect of articulatory suppression and no significant interaction between group and condition was found. These results provide no evidence to imply that suppressing inner speech has an effect on planning in adults (as some previous literature has suggested) and that adults with ASD are different in their use of inner speech or visual-spatial processing during planning when compared to TD adults. This pattern of results fails to fit with either the delayed or atypical development model. More research is needed which combines data from both adults and children to further distinguish between the delayed and atypical development models.

4. Post-hoc Analysis of Experiments 1 and 2

4.1 Introduction

Previous research suggests differences in planning strategies between TD individuals and individuals with ASD are seen regardless of whether the sample is made up of children (Holland & Low, 2010; Russell-Smith et al. 2014; Whitehouse et al., 2006), adolescents (Wallace et al. 2009) or adults (Williams et al. 2012). As Experiment 1 (using children between the ages of 7 and 12) and Experiment 2 (using adults) have yet to find any evidence of this, I conducted a post-hoc analysis which combined and compared the data from these two experiments to look for an interaction between group and condition with more power and to look for an interaction between age, group and condition.

4.1.1 Hypotheses. If the atypical development model is accurate then individuals with ASD will show a greater impairment in performance under visual-spatial suppression, a reduced impairment in performance under articulatory suppression and no difference in the effect of private speech on performance when compared to TD individuals. If the delayed development model is accurate then group differences will only be seen in the child sample who will also show a difference in the effect of private speech on planning (i.e. private speech will improve planning performance for children with ASD only).

4.2 Methods

The sample was made up of all participants from Experiments 1 and 2. This created a sample of 62 participants in total. The data analysis was identical to that of Experiment 1 except scores were standardised relative to the TD group for each condition for adults and children and for accuracy and time. As scores were standardised to the TD group for each condition and for both adults and children the chances of seeing a significant main effect of condition or age is reduced. No preliminary t-tests, or correlations were conducted.

4.3 Results

Two mixed-design ANOVAs, using group (two levels) as the between-subjects independent variable, condition (three levels – baseline, articulatory suppression and visual-spatial suppression) as the within-subjects independent variable and accuracy and time as the dependent variables, were conducted. Mauchly's test of sphericity was not significant for accuracy [$\chi2(2) = 4.73$, $p = .094$] but was for time [$\chi2(2) = 6.24$, $p = .044$, $\varepsilon = .906$] and so the Greenhouse-Giesser correction was applied for time. There was no significant main effect of condition for accuracy [$F(2, 116) = 1.46$, $p = .236$] or time [$F(2, 116) = 0.79$, $p = .446$], no significant main effect of group for accuracy [$F(1, 58) = 0.22$, $p = .640$] or time [$F(1, 58) = 0.10$, $p = .755$], no significant main effect of age for accuracy [$F(1, 58) = 0.11$, $p = .745$] or time [$F(1, 58) = 0.22$, $p = .645$], no significant interaction between condition and group for accuracy [$F(1, 116) = 1.46$, $p = .236$] or time [$F(2, 116) = 0.79$, $p = .446$], no significant interaction between condition and age for accuracy [$F(1, 116) = 2.22$, $p = .113$] or time [F(2, 116) =, 1.30 $p = .275$], no significant interaction between age and group for accuracy [$F(1, 58) = 0.11$, $p = .745$] or time [$F(1, 58) = 0.22$, $p = .645$] and no significant interaction between age, group and condition for accuracy [$F(1, 116) = 2.22$, $p = .113$] or time [$F(2, 116) = 1.30$, $p = .275$]. Descriptive statistics can be seen in Table 6.

Table 6

Descriptive Statistics (Articulatory and Visual-Spatial Suppression). Means and Standard Deviations for the baseline, articulatory suppression and visual-spatial suppression conditions in each diagnostic group for both Experiments 1and 2 combined (standardised to the TD group for both adults and children in each condition).

Measure	Condition	Group (N)	Mean	SD
Accuracy (Z-Score)	Baseline	ASD (26)	0.20	1.52
		TD (36)	0.00	1.00
	Articulatory Suppression	ASD (26)	-0.26	0.76
		TD (36)	0.00	1.00
	Visual-spatial Suppression	ASD (26)	-0.23	0.83
		TD (36)	0.00	1.00
Time (Z-Score)	Baseline	ASD (26)	-0.21	0.83
		TD (36)	0.00	1.00
	Articulatory Suppression	ASD (26)	0.05	1.43
		TD (36)	0.00	1.00
	Visual-spatial Suppression	ASD (26)	-0.10	0.84
		TD (36)	0.00	1.00

Two more mixed-design ANOVAs, using group (two levels) as the between-subjects independent variable, condition (two levels – baseline and private speech) as the within-subjects independent variable and accuracy and time as the dependent variables, were conducted. These ANOVAs excluded participants who had received a private speech score less than 2. Levene's test indicted the assumption of equal variances was violated for baseline accuracy [$F(3, 49) = 3.78$, $p = .016$] but not for private speech accuracy [$F(3, 49) = 0.92$, $p = .436$], baseline time [$F(3, 49) = 0.06$, $p = .980$] or private speech time [$F(3, 49) = 0.91$, $p = .444$]. There was no significant main effect of condition for accuracy [$F(1, 49) = 0.30$, $p = .585$] or time [$F(1, 49) = 0.11$, $p = .742$], no significant main effect of group for accuracy [$F(1, 49) = 0.05$, $p = .817$] or time [$F(1, 49) = 0.10$, $p = .753$], no significant main effect of age for accuracy [$F(1, 49) = 0.05$, $p = .829$] or time [$F(1, 49) = 0.45$, $p = .507$], no significant interaction between condition and group for accuracy [F(1, 49) = 0.30, p = .585] or time [F(1, 49) = 0.11, p = .742], no significant interaction between condition and age [F(1, 49) = 1.61, p = .210] or time [F(1, 49) = 0.30, p = .588], no significant interaction between age and group for accuracy [F(1, 49) = 0.05, p = .829] or time [F(1, 49) = 0.45, p = .507], and no significant interaction between age, group and condition for accuracy [F(1, 49) = 1.61, p = .210] or time [F(1, 49) = 0.30, p = .588]. Descriptive statistics can be seen in Table 7.

Table 7

Descriptive Statistics (Private Speech). Means and Standard Deviations for the baseline and private speech condition in each diagnostic group, after those who failed to engage in private speech were excluded, for both Experiments 1 and 2 combined (standardised to the TD group for both adults and children in each condition).

Measure	Condition	Group (N)	Mean	SD
Accuracy (Z-Score)	Baseline	ASD (20)	0.00	1.13
		TD (33)	0.00	1.00
	Private Speech	ASD (20)	-0.12	0.47
		TD (33)	0.00	1.00
Time (Z-Score)	Baseline	ASD (20)	-0.07	1.00
		TD (33)	0.00	1.00
	Private Speech	ASD (20)	-0.13	0.80
		TD (33)	0.00	1.00

4.4 Discussion

The atypical development model predicts that differences in the use of inner speech and visual-spatial processing between TD individuals and individuals with ASD will be consistent across the lifespan. Alternatively, the delayed development model predicts these differences will disappear in adulthood.

There was no significant main effect of group, age or condition (on accuracy or time) and no meaningful interactions between factors (for accuracy or time) for any of the four mixed-design ANOVAs. As scores were standardised relative to the TD group for each condition and for both adults and children the chance of seeing a significant main effect of condition is reduced and therefore no effect was found despite an effect of condition being found in both Experiments 1 and 2.

4.4.1 Previous Literature. As no reliable group differences were found, these results fail to fit with either the atypical or delayed development models and challenge much of the previous literature (Holland & Low, 2010; Wallace et al. 2009; Whitehouse et al., 2006; Williams et al. 2012). As in Experiments 1 and 2, sample size and a lack of power may be an issue as the groups are small. However, the sample size was much larger than those used in the previous literature when age as a factor was ignored. This means this analysis had more power than the previous literature, and both the previous experiments, to detect an interaction between group and condition if one existed and was constant across the lifespan as previous studies suggest (Holland & Low, 2010; Wallace et al. 2009; Whitehouse et al., 2006; Williams et al. 2012).

4.4.2 Conclusion. No significant effect of articulatory suppression, no significant interaction between group and condition and no significant interaction between age, group and condition was found. These results provide no evidence to suggest that suppressing inner speech affects planning ability, that individuals with ASD are different in their use of inner speech and visual-spatial processing during planning when compared to TD individuals, and that these two findings change across the lifespan. This pattern of results fails to fit with either the delayed or atypical development model and contradicts much of the previous literature.

5. General Discussion

Two experiments (one sampling children and another sampling adults) and a post-hoc analysis which combined the data of Experiments 1 and 2 found no significant group differences in the effects of experimentally manipulated private speech, articulatory suppression and visual-spatial suppression on planning between individuals with and without ASD. The pattern of results from this study do not fit with either the delayed or atypical development models and provide no evidence to suggest that individuals with ASD (both adults and children) show differences in their inner speech or visual-spatial processing use when engaging in goal-directed behaviour. This challenges the validity of a large body of previous literature (Holland & Low, 2010; Wallace et al. 2009; Whitehouse et al., 2006; Williams et al. 2012). No significant effects of articulatory suppression were seen across either group and no significant effects of dual tasks on accuracy were seen (significant effects of dual tasks were only seen on time). This challenges the validity of much of the literature on TD children above the age of 7 which has shown articulatory suppression impairs performance during goal-directed behaviour and has shown dual tasks to affect the number of moves in the Tower of London task (see Alderson-Day & Fernyhough, 2015 for a review). However, the results of Experiment 2 support some previous studies which provided no evidence to suggest TD adults rely on inner speech when planning (Gilhooly et al., 2002; Law et al., 2013; Miyake et al., 2004; Wynn et al., 1999). If the results of the current study are accurate, this would suggest much of the previous literature has been subject to a lack of power (when small sample sizes increase the likelihood of obtaining false significant results), confirmation bias (when the desire for a desired result causes an experimenter to subconsciously bias the methods or analysis in favour of finding this result) and publication bias (when journals do not publish failed replications, causing the finding to be viewed as more robust than it actually is).

5.1 Limitations

One major limitation with the current study is that the spatial tapping task may not be a valid form of visual-spatial suppression. While articulatory suppression is theoretically a valid way to supress inner speech and has been shown to by a large body of research (Cowan, et al., 1987; Ford & Silber, 1994; Hasselhorn & Grube, 2003; Hitch et al., 1989; Holland & Low, 2010; Lidstone et al., 2010; Russell-Smith et al., 2014; Whitehouse et al., 2006), tapping two points to suppress the visual-spatial processing used during the Tower of London task is arguably less theoretically valid and little research has used it in this context. It could be the case that the tapping task supresses the spatial aspect of visual-spatial processing but leaves the visual aspect needed to imagine potential solutions unaffected. This would mean any differences in the use of visual-spatial processing between the two groups would not be detected and could explain why this study and Holland and Low (2010) found no significant differences in the effect of visual-spatial suppression between the TD and ASD group. Evidence of dissociations between visual processing and spatial processing come from neuroimaging research (Courtney, Petit, Maisog, Ungerleider & Haxby, 1998) and impaired spatial processing but not visual processing in children with William's Syndrome (Vicari, Bellucci & Carlesimo, 2007).

Unlike articulatory suppression, visual-spatial suppression also requires the movement of both hands simultaneously which could potentially interfere with the task by making it harder for participants to complete the task physically (i.e. moving the blocks) as opposed to mentally (i.e. working out the correct solution). It was observed that many participants struggled to move blocks while tapping in time (either failing to tap in time or having to move blocks in time with the tapping task). This could be the reason visual-spatial suppression only affected time and not accuracy. This introduces a potential confound to the paradigm. Evidence has shown that individuals with ASD are especially impaired at the simultaneous execution of multiple arm and hand movements compared to individuals without ASD (McAuliffe, Pillai, Tiedemann, Mostofsky & Ewen, 2016). This would make it difficult to conclude if any group differences seen were due to differences in the use of visual-spatial processing or due to differences in the ability to cope with simultaneous movements. However, no evidence of group differences were found meaning this is unlikely to have affected the results.

A possible issue is that the planning measure may not have been sensitive enough to detect effects on planning ability. This study found no significant effects for planning accuracy. Furthermore, participants in both experiments were incredibly accurate, with children taking (on average) between 0.57 and 1.07 moves over the minimum and adults taking between 0.57 and 0.99 moves over the minimum depending on the group and condition. Williams et al (2012) used harder trials in their study on adults (ranging from 5 to 13 moves) and found accuracy to be poorer (averaging between 0.83 and 1.47 moves over the minimum depending on the group and condition). This may imply that performance was at ceiling and that harder trials were needed to detect effects of dual tasks on planning accuracy. The only studies to assess the effects of articulatory suppression on planning ability between children with ASD and TD children (Holland & Low, 2010; Wallace et al., 2009) did not use the Five-Disk Tower of London task (they used the Tower of Hanoi and the original Tower of London task respectively) and did not report the trials used so the difficulty of these trials cannot be compared with the difficulty of Experiment 1. While, Ward and Allport (1997) found that the number of overall moves and the number of sub-goal moves significantly predicted time taken and errors made, it was also found that the number of sub-goal chunks (i.e. collections of consecutive sub-goal moves) was a better predictor of error rate and that the number of disks that could be moved at the start of each sub-goal chunk also affected difficulty. These predictors of difficulty were not considered during the creation of the trials used in the current study and could have caused performance to be at ceiling for accuracy, leading to a lack of significant effects for accuracy.

Another potential limitation with the study is biases in sampling. For example, several children with ASD from Experiment 1 were excluded from the entire experiment for being unable to do the tasks and several were excluded from the private speech condition for not engaging in task focused private speech. This means children with ASD who are low-functioning (i.e. more intellectually and socially impaired) are not represented in the sample and so it is difficult to generalise the results to the entire ASD population. This could lead to a reduction in group differences. Also, the participants from Experiment 2 were sampled through invitation and advertisement and had to travel to the University of Bristol to be tested which could have biased the sample to be mostly made up of adults with ASD who are high-functioning. This bias in the sample is highlighted by the fact that there were no significant differences in intelligence between the two groups in both Experiments 1 and 2 ($p < .700$, $d = 0.14$ and $p = .460$, $d = 0.45$ respectively) and the fact that in both experiments and the post-hoc analysis there was no significant main effect of group on planning ($p \geq .491$). Previous experiments have used samples made up of entirely high-functioning ASD samples and found group differences (Russell-Smith et al., 2014; Williams et al., 2012), suggesting this is not the reason for the disparity. Furthermore, strong significant differences between the ASD and TD groups on measures of ASD traits were found for both Experiments 1 and 2 ($p < .001$, $d = 2.35$ and $p < .001$, $d = 3.61$) confirming that the lack of a significant group/condition interaction is not due to the ASD group containing many individuals with mild symptom severity or the TD group containing individuals with undiagnosed ASD. However, 8 of the 11 ASD participants in Experiment 2 identified themselves as having been diagnosed with

Asperger's Syndrome, a subset of ASD characterised by less severe language and intellectual deficits (World Health Organization, 1992). Individuals with Asperger's might develop more typical private/inner speech use than individuals with standard ASD which could lead to more typical development. This again means it is difficult to generalise the results to the entirety of the ASD population and that group differences could have been reduced.

Another sampling bias is that the ASD group in Experiment 1 is mostly made up of males which reflects the real-world gender bias in ASD (Fombonne, 2002; Levy, Mandell, & Schultz, 2009; Windham et al., 2011), whereas the ASD group in Experiment 2 is more even (in fact having slightly more females). As ASD is believed to manifest differently in females who are shown to be better at 'masking' their social deficits (Ehlers & Gillberg, 1993; Goldman, 2013; Mattila et al., 2007) it could be the case that this would lead to more typical development and a reduction in group differences. Another issue is differences in gender between the groups. while it is not the case for Experiment 2, Experiment 1 (and many previous studies in this field) had a more even gender ratio in the control group than in the ASD group. Gender differences in visual-spatial processing abilities have been widely reported (Flaherty, 2005; Jansen & Heil, 2010), although whether these differences are real and whether they are innate or caused by social influence is a topic of much debate (Lippa, Collaer & Peters, 2010; Neuburger, Quaiser-Pohl & Jansen, 2012). If any differences in visual-spatial processing ability between males and females do exist (regardless of their cause) this would confound the results. This means that if individuals with ASD were found to be more reliant on visual-spatial processing than inner speech this could simply be due to the ASD group containing far more males and not due to ASD itself. While Wallace et al. (2009) and Whitehouse et al. (2006) had equal or similar gender ratios in both groups, Williams et al. (2012) and Holland and Low (2010) did not report gender information.

However, no evidence of group differences were found in this study meaning this is unlikely to have affected the results.

5.2 Future Research

As the replicability of the previous literature has been called into question and most of the experiments in this field have relatively small sample sizes (less than 30 in each group), a large-scale replication is necessary to assess whether the previous findings are valid or due to a lack of power, confirmation bias and publication bias. Future research also needs to focus on how planning strategies may vary within the autism spectrum. I propose a study which not just compares individuals with ASD and TD controls but individuals with different ASD profiles (i.e. split by gender, diagnosis and intellectual ability). However, this would require a more time-consuming data collection process which would include a full IQ test and testing many participants. This would also be difficult as it would require the planning task to be easy enough for individuals with severe intellectual impairments and hard enough for individuals without. I suggest a more sophisticated and realistic version of the computerised Tower of London task is used. Many of the children who were unable to complete the Tower of London task (and thus were excluded) kept forgetting the basic rules and would often try to move blocks from under other blocks or place blocks in a floating position despite regular reminders. A non-computerised version of the task would not allow them to attempt these actions, however, this would take longer to run and is more susceptible to human error. Alternatively, a more sophisticated computerised version of the task, which was 3D and based in a real life setting (i.e. on a table instead of floating in white space), which also let the children drag the blocks off each peg and place them on to the new peg instead of the blocks jumping from one position to another when the new position is tapped, may have helped the more intellectually impaired children grasp the basic concepts of the tasks. This study may also need to have a larger range of trials in terms of difficulty (varying the number of overall moves, the number of sub-goal chunks and the number of moveable disks at the start of each sub-goal chunk) to address concerns that performance was at ceiling for accuracy in the current study. Currently no studies in this field have compared different ASD profiles but identifying differences in planning strategies across the autism spectrum would increase our understanding of ASD and help us create planning strategies that would be tailored to each individual case of ASD and not just ASD in general. Future research should also replicate the current study with a more valid visual-spatial suppression task. For example, participants could be shown a picture before each trial and must complete a memory task afterwards (e.g. choosing the correct picture out of several very similar options or having to draw it). This arguably would be a far more valid method of visual-spatial suppression as it would require the participant to visualise something other than the planning task throughout the trial and does not load motor skills as it requires no movement during the planning task itself. A more analogous version of the articulatory suppression task could be used in conjunction with this where the participant must remember a list of words and either repeat the list after the trial or choose the correct list from several options (e.g. the same words in a different order).

5.3 Conclusion

To conclude, the results of this study call the replicability and validity of the previous finding that individuals with ASD are unimpaired by articulatory suppression during goal-directed behaviour (unlike TD controls) into question and suggest it may be a product of confirmation bias, publication bias and underpowered studies. However, the current study was also affected by several limitations that may have affected the results (i.e. possible biases in sampling, a planning measure which may have lacked sensitivity and a potentially flawed visual-spatial suppression paradigm). I suggest a large-scale replication should be carried out to address the questions raised by this study, which splits the ASD group into large subgroups (based on gender, diagnosis and intellectual ability), uses a more realistic computerised version of the Tower of London task, uses a larger range of trial difficulty and uses a more theoretically valid measure of visual-spatial suppression (such as holding an image in memory throughout a trial) to control for the limitations of the current study.

References

Alderson-Day, B. & Fernyhough, C. (2015). Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology. *Psychological Bulletin, 141*(5), 931-965.

Al-Namlah, A. S., Fernyhough, C., & Meins, E. (2006). Sociocultural influences on the development of verbal mediation: Private speech and phonological recoding in Saudi Arabian and British samples. *Developmental Psychology, 42,* 117–131.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR.* (4th ed., text revision.. ed.). Washington, D.C.: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence for task switching. *Journal of Experimental Psychology: General, 130*, 641–657

Barnard-Brak, L., Brewer, A., Chesnut, S., Richman, D., & Schaeffer, A. (2016). The sensitivity and specificity of the social communication questionnaire for autism spectrum with respect to age. *Autism Research, 9*(8), 838-845.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or highfunctioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1430), 361-374.

Baron-Cohen, S., & Swettenham, J. (1997). Theory of mind in autism: Its relationship to function and central coherence. *Handbook of autism and pervasive developmental disorders*, 880-893.

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *The British Journal of Psychiatry, 175*(5), 444–451.

Campbell, C., Landry, O., Russo, N., Flores, H., J., Jacques, S. & Burack, J. A. (2013). Cognitive Flexibility among Individuals with Down Syndrome: Assessing the Influence of Verbal and Nonverbal Abilities. *American Journal on Intellectual and Developmental Disabilities,118*(3), 193-200.

Campbell, C. A., Russo, N., Landry, O., Jankowska, A. M., Stubbert, E., Jacques, S. & Burack, J. A. (2017). Nonverbal, rather than verbal, functioning may predict cognitive flexibility among persons with autism spectrum disorder: A preliminary study. *Research in Autism Spectrum Disorders*, *38*, 19.

Constable, P. A., Ring, M., Gaigg, S. B. & Bowler, D. M. (2017). Problem-solving styles in autism spectrum disorder and the development of higher cognitive functions. *Autism*

Corsello, C., Hus, V., Pickles, A., Risi, S., Cook, E. H., Leventhal, B. L. & Lord, C. (2007). Between a ROC and a hard place: decision making and making decisions about using the SCQ. *Journal of Child Psychology and Psychiatry*, *48*(9), 932– 940.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G. & Haxby, J. V. (1998) An area specialized for spatial working memory in human frontalcortex. *Science, 279*, 347–51.

Cowan, N., Cartwright, C., Winterowd, C., & Sherk, M. (1987). An adult model of preschool children's speech memory. *Memory & Cognition, 15*(6), 511-517.

Dunn, D.M., Dunn, L.M. & Styles, B. (2009). *The British Picture Vocabulary Scale (3rd Ed. BPVS3).* London: GL Assessment.

Eaves, L. C., Wingert, H., Ho, H. H., & Mickelson, E. (2006). Screening for autism spectrum disorders with the Social Communication Questionnaire. *Journal of Developmental and Behavioral Pediatrics, 27*(2)

Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger syndrome. A total population study. *Journal of Child Psychology and Psychiatry, 34*(8), 1327–1350.

Fabricius, W. V. (1988). The development of forward search planning in preschoolers. *Child Development, 59*, 1473–1488.

Fagot, B., & Gauvain, M. (1997). Mother-child problem solving: Continuity through the early childhood years. *Developmental Psychology, 33*(3), 480-8.

Fernyhough, C. (1996). The dialogic mind: A dialogic approach to the higher mental functions. *New Ideas in Psychology, 14*(1), 47-62.

Flaherty, M. (2005). Gender differences in mental rotation ability in three cultures: Ireland, ecuador and japan. *Psychologia -An International Journal of Psychology in the Orient, 48*(1), 31-38.

Fombonne, E. (2002). Epidemiological trends in rates of autism. *Molecular Psychiatry, 7*(Suppl. 2), S4–S6.

Ford, S., & Silber, K. P. (1994). Working-memory in children—A developmental-approach to the phonological coding of pictorial material. *British Journal of Developmental Psychology, 12*, 165–175.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition*, *50*(1-3), 115-132.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39*, 3–27.

Gilhooly, K. J. (2005). Working memory and strategies in reasoning. In E. Newton & M. Roberts (Eds.), *Methods of thought: Individual differences in reasoning strategies* (pp. 57–80). Hove, UK: Psychology Press.

Gilhooly, K. J., Wynn, V., Phillips, L. H., Logie, R., & Della Sala, S. (2002). Visuo-spatial and verbal working memory in the Five-disc Tower of London task: An individual differences approach. *Thinking & Reasoning, 8*, 165–178.

Goldman, S. (2013). Opinion: Sex, gender and the diagnosis of autism—A biosocial view of the male preponderance. *Research in Autism Spectrum Disorders, 7*(6), 675–679.

Grandin, T. (1995). How people with autism think. In E. Schopler & G. B. Mesibov (Eds.), *Learning and cognition in autism* (pp. 137–156). New York, NY: Springer.

Halliday, M. S., Hitch, G. J., Lennon, B., & Pettipher, C. (1990). Verbal short-term memory in children: The role of the articulatory loop. *European Journal of Cognitive Psychology, 2*, 23–38.

Harvey, A. J., & Beaman, C. P. (2007). Input and output modality effects in immediate serial recall. *Memory, 15,* 693-700.

Hasselhorn, M., & Grube, D. (2003). The phonological similarity effect on memory span in children: Does it depend on age, speech rate, and articulatory suppression? *International Journal of Behavioral Development, 27,* 145–152.

Hirata, Shogo, Okuzumi, Hideyuki, Kokubun & Mitsuru (2016). Brief report: Effect of verbalization to other individuals on cognitive planning in children with autism spectrum disorders. *Postepy Psychiatrii i Neurologii, 25*(4), 231

Hitch, G. J., Woodin, M. E., & Baker, S. (1989). Visual and phonological components of working memory in children. *Memory & Cognition, 17*, 175–185.

Holland, L., & Low, J. (2010). Do children with autism use inner speech and visuo-spatial resources for the service of executive control? Evidence from suppression in dual tasks. *British Journal of Developmental Psychology, 28,* 369–391.

Hurlburt, R. T., Happe, F., & Frith, U. (1994). Sampling the form of inner experience in 3 adults with Asperger syndrome. *Psychological Medicine, 24,* 385–395.

Jansen, P., & Heil, M. (2010). Gender differences in mental rotation across adulthood. *Experimental Aging Research, 36*(1), 94-104.

Jarrold, C., & Citroën, R. (2013). Reevaluating key evidence for the development of rehearsal: Phonological similarity effects in children are subject to proportional scaling artifacts. *Developmental Psychology, 49*, 837–847

Joseph, R. M., Steele, S. D., Meyer, E., & Tager-Flusberg, H. (2005). Selfordered pointing n children with autism: Failure to use verbal mediation in the service of working memory? *Neuropsychologia, 43,* 1400–1411.

Kalandadze, T., Norbury, C., Nærland, T., & Næss, K. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism: The International Journal of Research and Practice, 22*(2), 99-117.

Kray, J., Eber, J., & Karbach, J. (2008). Verbal self-instructions in task switching: A compensatory tool for action control deficits in childhood and old age? *Developmental Science, 11*, 223–236

Law, A. S., Trawley, S. L., Brown, L. A., Stephens, A. N., & Logie, R. H. (2013). The impact of working memory load on task execution and online plan adjustment during multitasking in a virtual environment. *The Quarterly Journal of Experimental Psychology, 66*, 1241–1258.

Levy, S. E., Mandell, D. S., & Schultz, R. T. (2009). Autism. *Lancet, 374*, 1627–1638.

Lidstone, J. S., Fernyhough, C., Meins, E., & Whitehouse, A. J. O. (2009). Brief report: Inner speech impairment in children with autism is associated with greater nonverbal than verbal skills. *Journal of Autism and Developmental Disorders, 39*, 1222–1225.

Lidstone, J. S. M., Meins, E., & Fernyhough, C. (2010). The roles of private speech and inner speech in planning during middle childhood: Evidence from a dual task paradigm. *Journal of Experimental Child Psychology, 107*, 438–451.

Lidstone, J. S., Meins, E., & Fernyhough, C. (2012). Verbal mediation of cognition in children with specific language impairment. *Development and Psychopathology*, *24*, 651–660.

Lippa, R., Collaer, M., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior, 39*(4), 990-7.

Mattila, M., Kielinen, M., Jussila, K., Linna, S., Bloigu, R., Ebeling, H., & Moilanen, I. (2007). An epidemiological and diagnostic study of Asperger syndrome according to four sets of diagnostic criteria. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(5), 636–646.

McAuliffe, D., Pillai, A., Tiedemann, A., Mostofsky, S., & Ewen, J. (2017). Dyspraxia in asd: Impaired coordination of movement elements. *Autism Research, 10*(4), 648-652.

Mitsuhashi, S., Hirata, S., & Okuzumi, H. (2018). Role of Inner Speech on Serial Recall in Children with ASD: A Pilot Study Using the Luria Hand Test. *Autism research and treatment*, *2018*.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*, 123–142.

Monsell, S., & Driver, J. (Eds.). (2000). *Attention and performance: Vol. 18. Control of cognitive processes.* Cambridge, MA: MIT Press.

Murray, D. J. (1966). Vocalization-at-presentation and immediate recall, with varying recall methods. *Quarterly Journal of Experimental Psychology, 18,* 9-18.

Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 523–528.

Neuburger, S., Quaiser-Pohl, C., Jansen, P., & Heil, M. (2012). A threat in the classroom: Gender stereotype activation and mental-rotation performance in elementary-school children. *Zeitschrift Fur Psychologie / Journal of Psychology, 220*(2), 61-69.

Ozonoff, S. (1995) 'Reliability and Validity of the Wisconsin Card Sorting Test in Studies of Autism', *Neuropsychology 9*, 491–500.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive Function Deficits in High-Functioning Autistic Individuals: Relationship to Theory of Mind. *Journal of Child Psychology and Psychiatry, 32*(7), 1081-1105.

Ozonoff, S., South, M., & Miller, J. (2000). DSM-IV-Defined Asperger Syndrome: Cognitive, Behavioral and Early History Differentiation from High-Functioning Autism. *Autism, 4*(1), 29-46.

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology, 5,* 176–206.

Peirce, J.W. (2007) PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2):8-13

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51-87.

Phillips, L. H., Wynn, V., Gilhooly, K. J., Della Sala, S., & Logie, R. H. (1999). The role of memory in the Five-disc Tower of London task. *Memory, 7*, 209–231.

Piaget, J. (1959). *The language and thought of the child*. Hove, UK: Psychology Press

Raven, J., Raven, J.C., & Court, J.H. (2003) *Manual for Raven's Progressive Matrices and Vocabulary Scales.* San Antonio, TX: Harcourt Assessment.

Russell, J., Jarrold, C., & Hood, B. (1999). Two intact executive capacities in children with autism: Implications for the core executive dysfunctions in the disorder. *Journal of Autism and Developmental Disorders, 29,* 103–112.

Russell-Smith, S. N., Comerford, B. J. E., Maybery, M. T. & Whitehouse, A. J. O. (2014). Brief Report: Further Evidence for a Link Between Inner Speech Limitations and Executive Function in High-Functioning Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 44*(5), 1236

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire manual.* Los Angeles, CA: Western Psychological Services.

Sayhoun, C. P., Beliveau, J. W., Soulieres, I., Schwartz, S., & Moody, M. (2010). Neuroimaging of the functional and structural networks underlying visuospatial vs. linguistic reasoning in high-functioning autism. *Neuropsychologia, 48,* 86–95.

Schuh, J. M., Eigsti, I. (2012). Working Memory, Language Skills, and Autism Symptomatology. *Behavioral Sciences, 2*(4), p. 207

Snow, A. N. & Lecavalier, L. (2008). Sensitivity and specificity of the modified checklist for autism in toddlers and the social communication questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism, 12*(6), 627–644.

Soulières, I., Dawson, M., Samson, F., Barbeau, E., Sahyoun, C., Strangman, G., . . . Mottron, L. (2009). Enhanced visual processing contributes to matrix reasoning in autism. *Human Brain Mapping, 30*(12), 4082-4107.

Soulières, I., Zeffiro, T. A., Girard, M. L., & Mottron, L. (2011). Enhanced mental image mapping in autism. *Neuropsychologia*, *49*(5), 848-857.

Spain, D., Sin, J., Linder, K., McMahon, J., & Happé, F. (2018). Social anxiety in autism spectrum disorder: A systematic review. *Research in Autism Spectrum Disorders, 52*, 51-68.

Vicari, S., Bellucci, S., & Carlesimo, G. (2007). Visual and spatial working memory dissociation: Evidence from williams syndrome. *Developmental Medicine & Child Neurology, 45*(4), 269-273.

Vygotsky, L. S. (1934/1987). *Thinking and speech. The collected works of Lev Vygotsky (Vol. 1).* New York: Plenum Press. (Original work published 1934)

Wallace, G. L., Silvers, J. A., Martin, A., & Kenworthy, L. E. (2009). Brief report: Further evidence for inner speech deficits in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 1735–1739.

Wang, X., Logie, R. H., & Jarrold, C. (2016). Interpreting potential markers of storage and rehearsal: Implications for studies of verbal short-term memory and neuropsychological cases. *Memory & Cognition*, 1-12.

Ward, G., & Allport, A. (1997). Planning and Problem solving Using the Five disc Five-disc Tower of London Task. *Quarterly Journal of Experimental Psychology, 50*(1), 49-78.

Wechsler, D. (2014). *Wechsler Intelligence Scale for children (5th ed., WISC-VCDN)*. Toronto, Ontario: Pearson Canada Assessment.

Whitehouse, A. J. O., Maybery, M. T., & Durkin, K. (2006). Inner speech impairments in autism. *Journal of Child Psychology and Psychiatry, 47,* 857–865.

Williams, D., Bowler, D., & Jarrold, C. (2012). Inner speech is used to mediate short-term memory, but not planning, among intellectually high-functioning adults with autism spectrum disorder. *Development and Psychopathology, 24*(1), 225-239.

Williams, D., Happe, F., & Jarrold, C. (2008). Intact inner speech use in autism spectrum disorder: Evidence from a short-term memory task. *Journal of Child Psychology and Psychiatry, 49,* 51–58.

Windham, G. C., Anderson, M. C., Croen, L. A., Smith, K. S., Collins, J., & Grether, J. K. (2011). Birth prevalence of autism spectrum disorders in the San Francisco Bay Area by Demographic and Ascertainment Source Characteristics. *Journal of Autism and Developmental Disorders, 41*, 1362–1372.

Winsler, A., Abar, B., Feder, M. A., Schunn, C. D., & Rubio, D. A. (2007). Private speech and executive functioning among high-functioning children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 37,* 1617–1635.

Winsler, A., Diaz, R. M., & Montero, I. (1997). The role of private speech in the transition from collaborative to independent task performance in young children. *Early Childhood Research Quarterly, 12*, 59–79.

Winsler, A., & Naglieri, J. (2003). Overt and covert verbal problem- solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5 to 17. *Child Development, 74*, 659–678.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S. & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ: a preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders. 35*(3), 331–5.

World Health Organization. (1992). *The ICD-10 classification of* ***mental*** *and behavioural disorders: Clinical descriptions and diagnostic guidelines.* Geneva: World Health Organization.

Appendix A

The Computerised Tower of London Task

The experimenter inputted the participant number, the condition, the age (which determined the difficulty) and selected the list of trials on the main menu (see Figure A1). Before each list, four pages of instructions were presented (see Figure A2 - A5). The participant pressed a "Next" button at the bottom of the first three pages of instructions to move to the next page. On the final page there was a "START" button to start the task. Once the trial had started the participant had to type the number of moves they thought it would take into a box under the puzzle (right of the label which said "How many moves?") and press the "START" button to the right of this before they could start moving blocks (see Figure A6). If they pressed start before typing in a number a red arrow appeared pointing to the box where they had to type the number from above (See Figure A7). After pressing the "START" button with an estimate typed into the box the button (and the arrow) disappeared. To the right of this a counter (next to the label which said "Moves:") counted the number of moves the participant had taken. If the participant tried to move a block that could not be moved or tried to move a block to an invalid position the Windows "default beep" sound was played to make the participant aware they were attempting an invalid move. Once the participant had clicked on a block they wished to move, a shadow appeared along its bottom and right edge highlighting it. Once the blocks were all in place the word "Correct!" appeared under the puzzle in green regardless of whether they did it in the correct number of moves, guessed this number correctly or correctly guessed how many moves it would take them (See Figure A8). After completing each trial, the participant was taken to a break screen containing a button which said "Next" (See Figure A9). Pressing this started the next trial. At the end of the list the participants were taken to a screen with a button that said "Finish" (See Figure A10). Pressing this button returned them to the main menu.

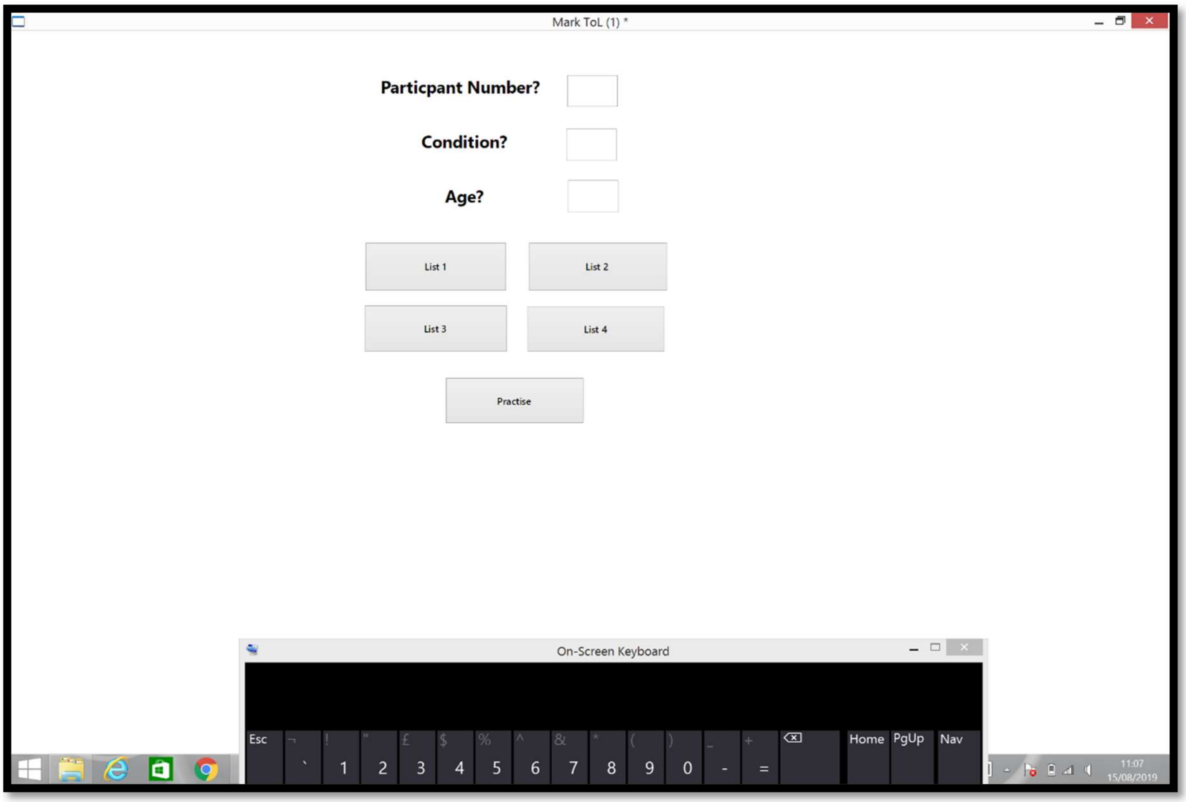

Figure A1. Screenshot of computerised Tower of London task main menu.

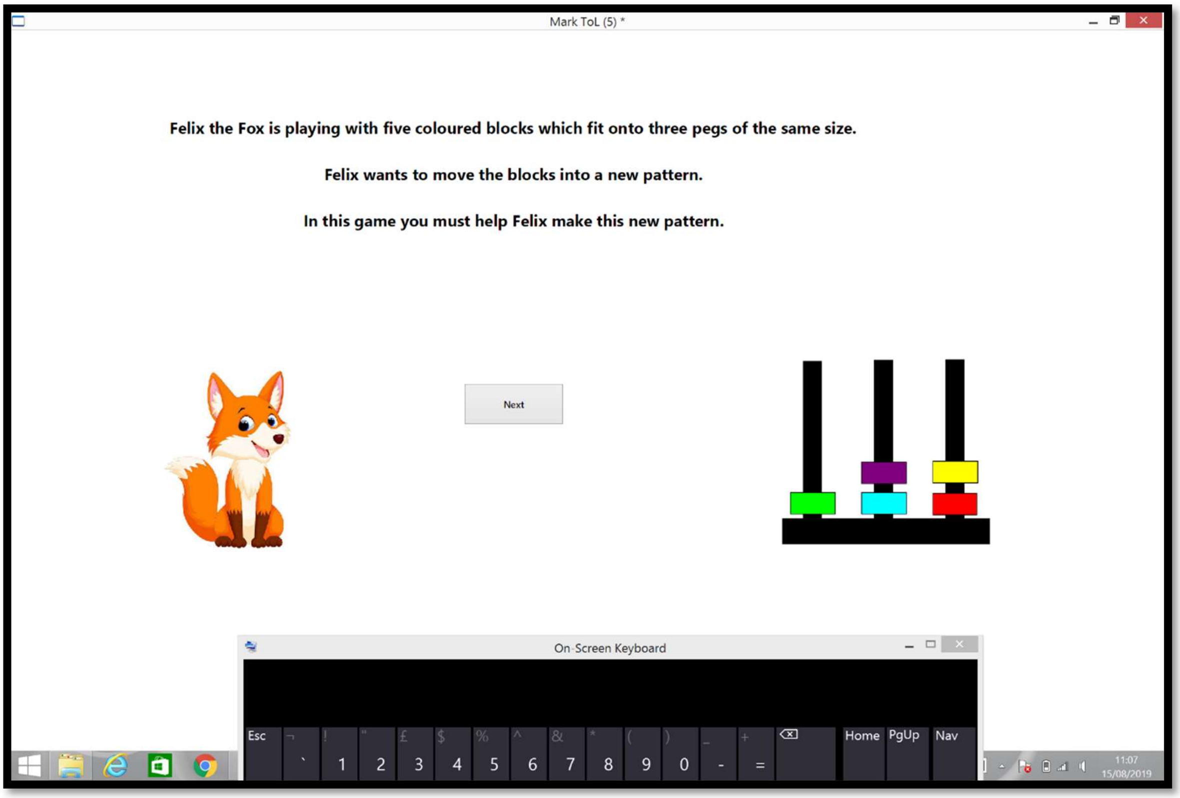

Figure A2. Screenshot of computerised Tower of London task instruction page 1.

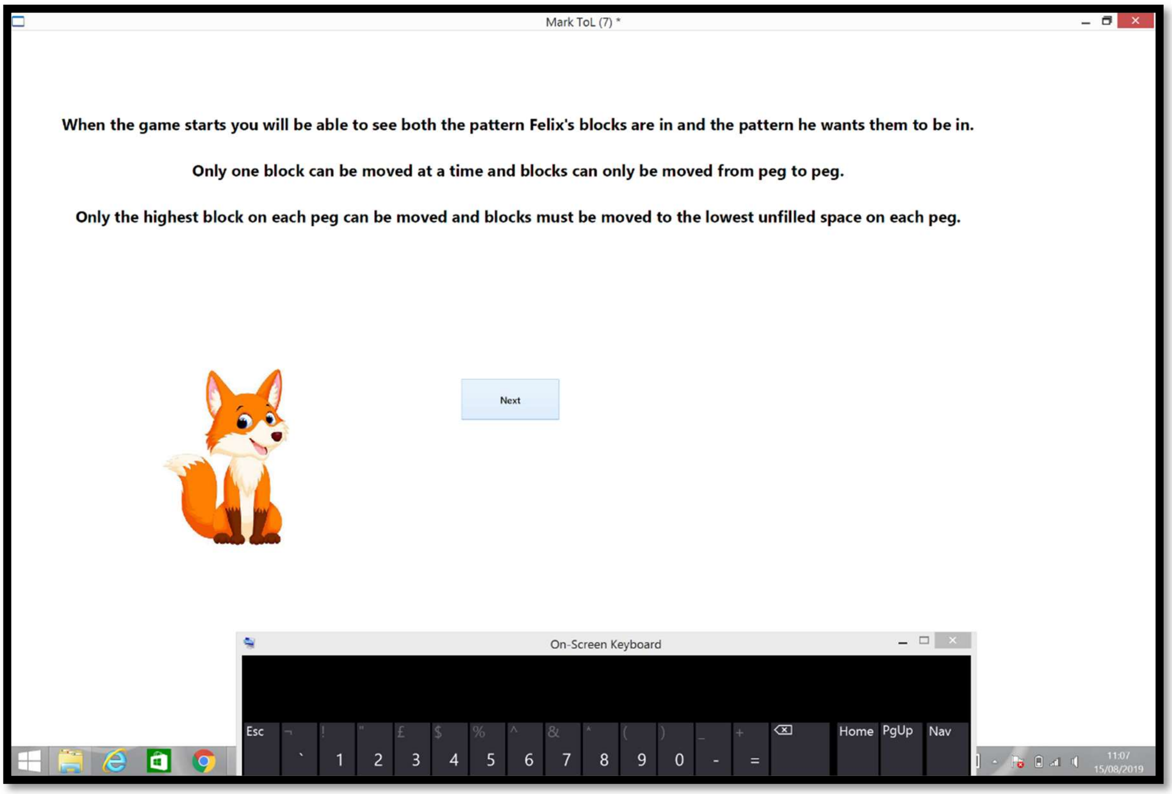

Figure A3. Screenshot of computerised Tower of London task instruction page 2.

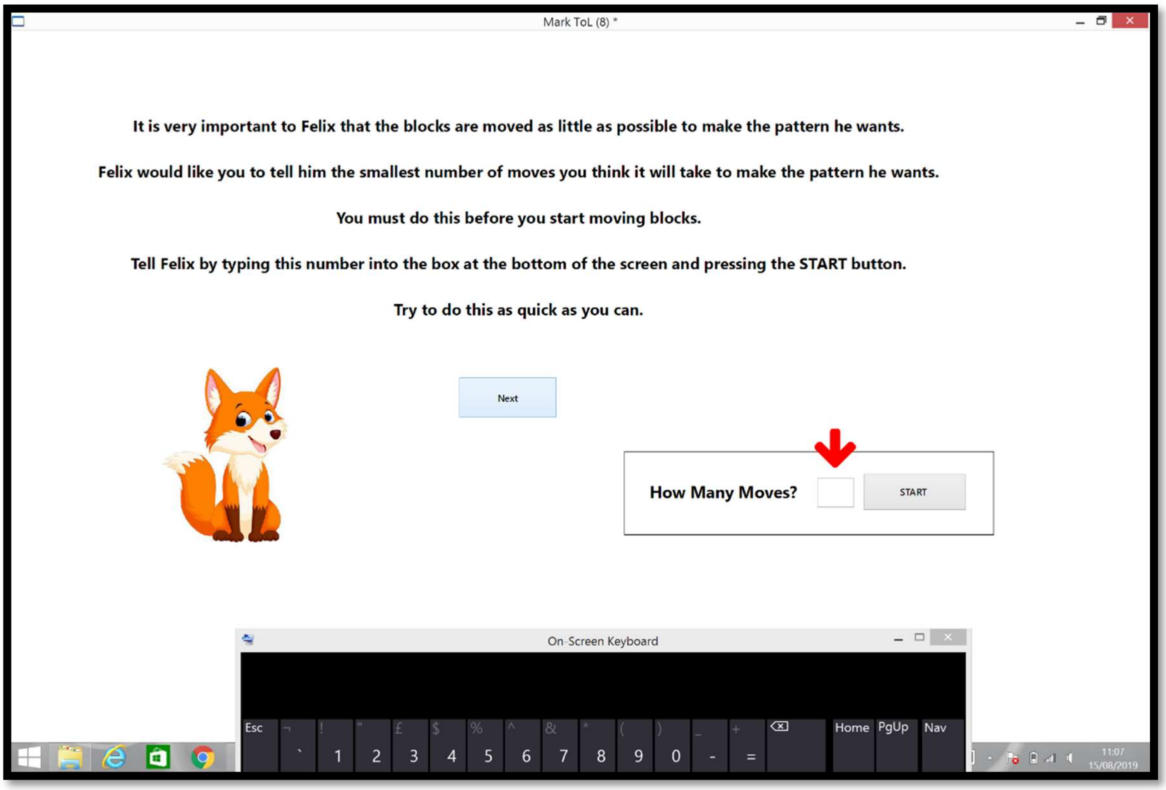

Figure A4. Screenshot of computerised Tower of London task instruction page 3.

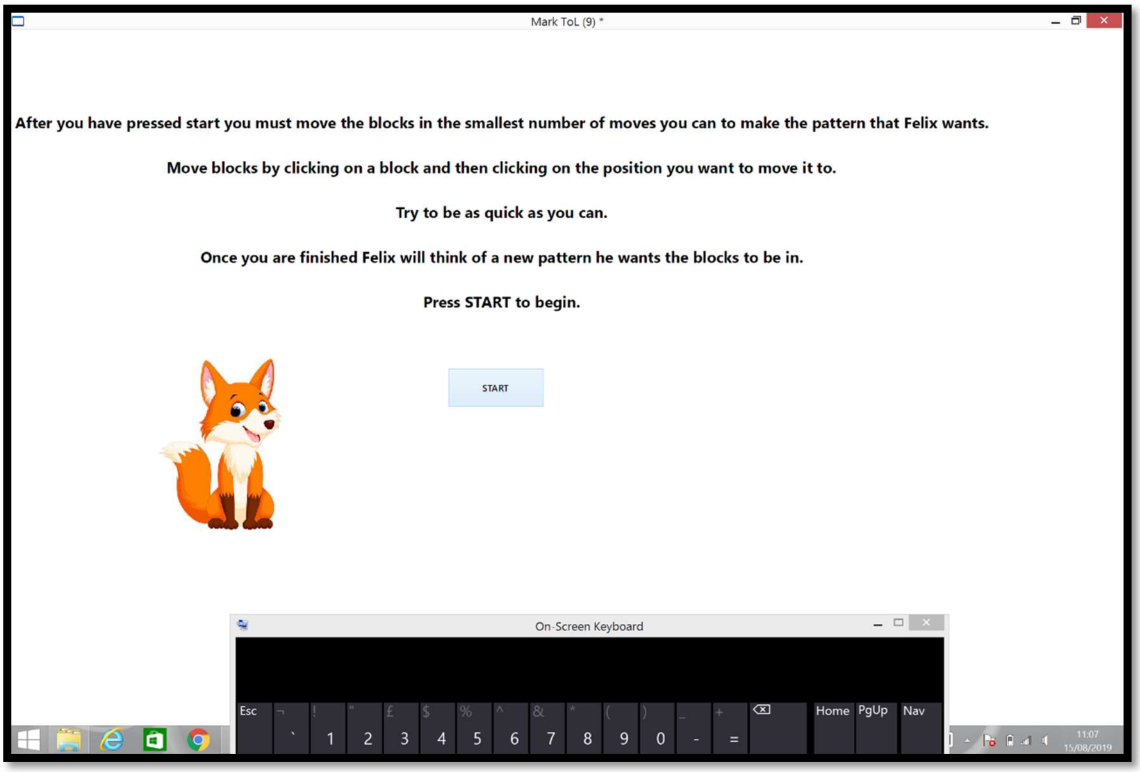

Figure A5. Screenshot of computerised Tower of London task instruction page 4.

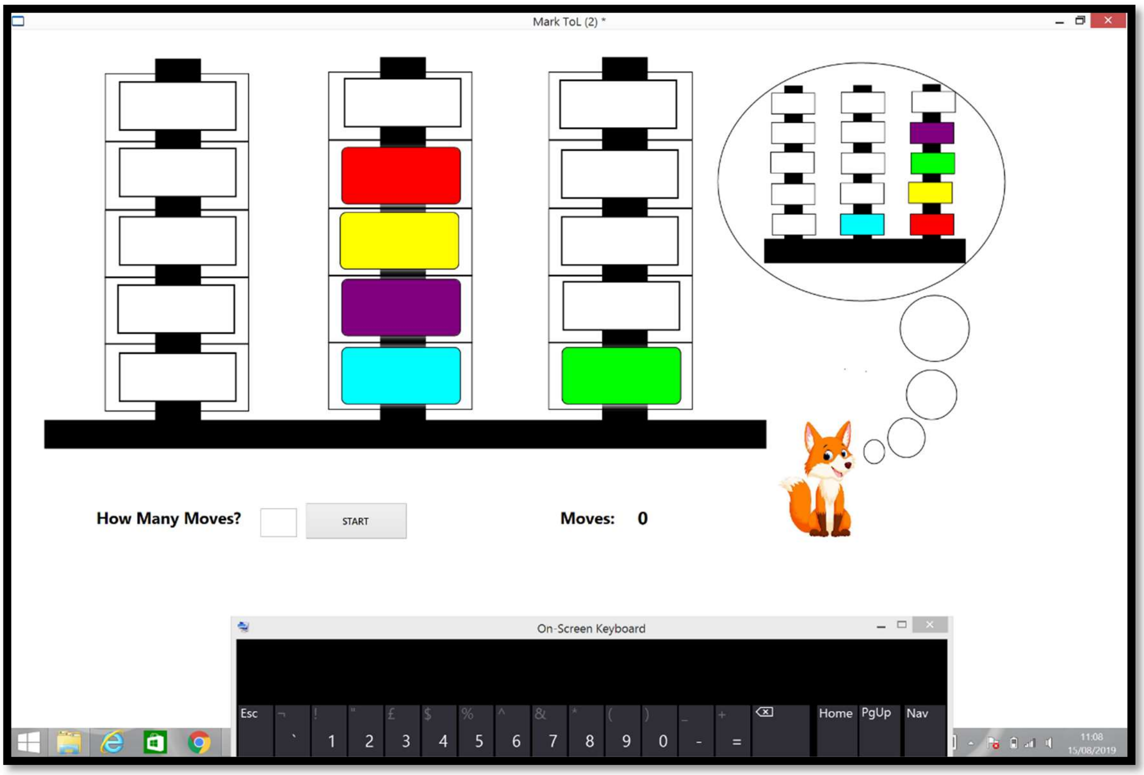

Figure A6. Screenshot of computerised Tower of London task trial before estimate.

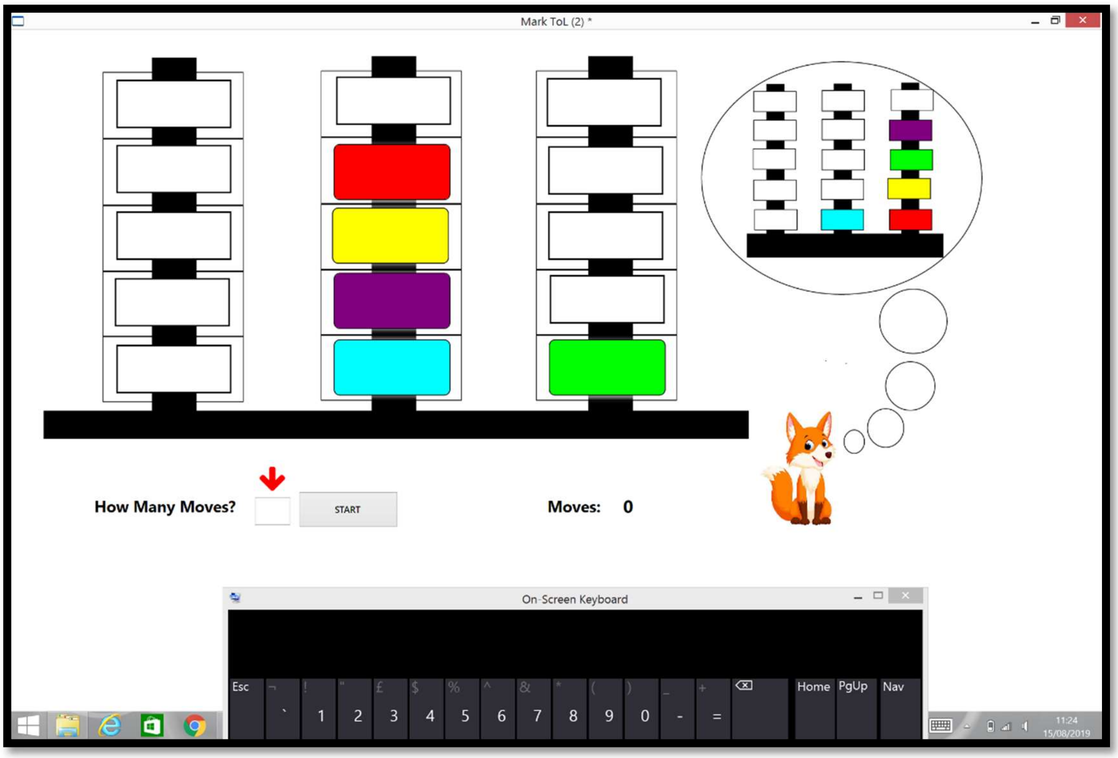

Figure A7. Screenshot of computerised Tower of London task trial if "START" button is pressed without an estimate being made.

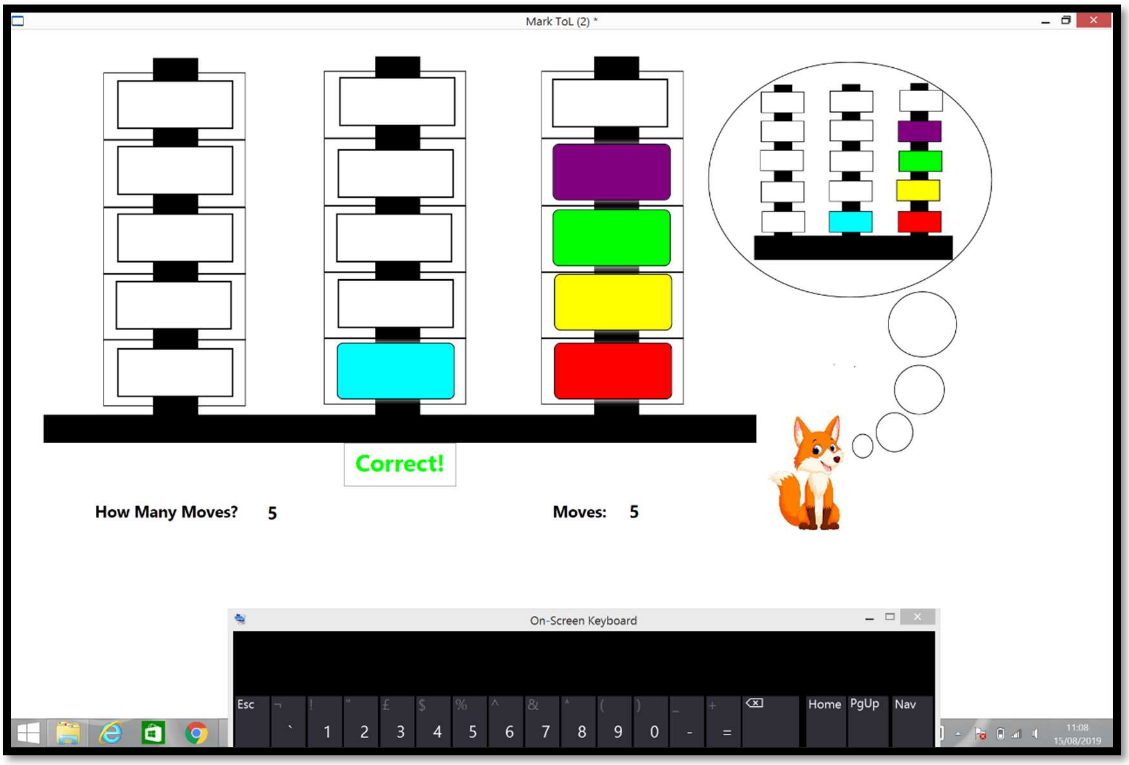

Figure A8. Screenshot of computerised Tower of London task trial at completion.

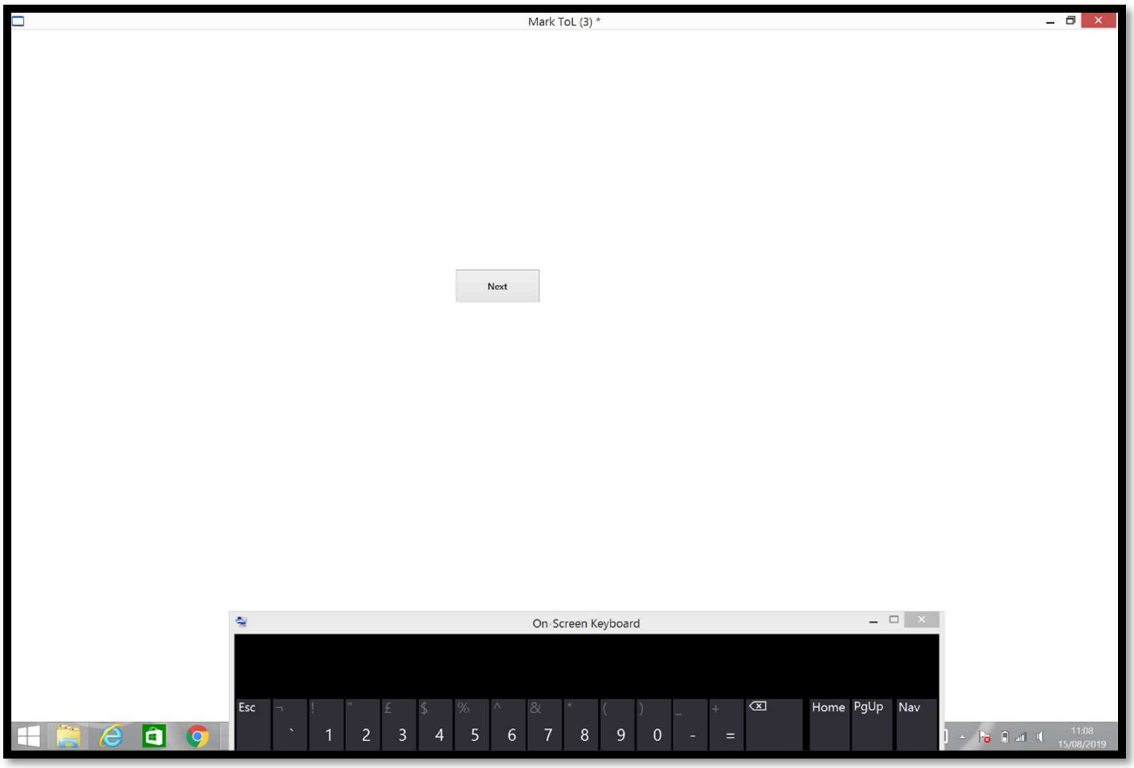

Figure A9. Screenshot of computerised Tower of London task break screen.

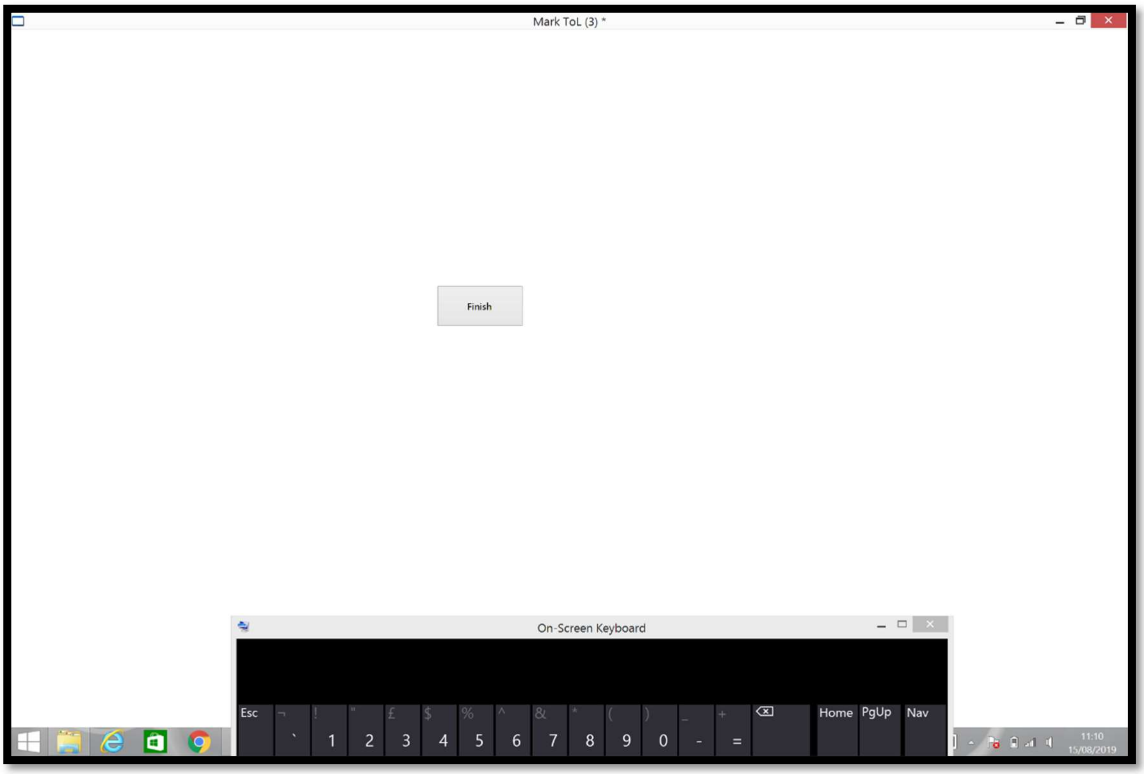

Figure A10. Screenshot of computerised Tower of London task end screen.

Appendix B

Trials for Experiments 1 and 2

Table B1 shows the trials of each list in the order they were presented, along with their difficulty and the location of the blocks at the start and the end (given as two lists of five numbers each between 1 and 15) for Experiments 1 and 2. The first number describes the location of the red block, the second number describes the location of the yellow block, the third number describes the location of the purple block, the fourth number describes the location of the blue block and the fifth number describes the location of the green block. Each peg has five available spaces on it. The spaces on the first peg are numbered 1-5, the spaces on the second peg are numbered 6-10 and the spaces on the third peg are numbered 11-15, with the lowest number being the lowest space and the highest number being the highest space. Trials of difficulty level 6.2, 7.2 and 7.3 (in bold) were omitted for children (Experiment 1) and trials of difficulty level 2.0 (italicised) were omitted for adults (Experiment 2).

Table B1

Order of trials in each list, their difficulty and the start and end configurations of their blocks for Experiments 1 and 2. Trials in bold were omitted for Experiment 1 and trials in italics were omitted form experiment 2.

List	Start Positions	End Positions	Difficulty
Practise	1,6,7,8,2	1,6,7,8,9	1.0
	1,2,3,6,11	1,2,7,6,3	2.0
	1,6,7,11,12	12,6,7,11,1	3.1
1	1,2,6,7,11	6,13,14,12,11	4.0
	1,2,3,6,11	**13,12,15,14,11**	**6.2**
	1,6,11,12,13	**6,11,9,8,7**	**7.2**
	9,8,7,6,11	11,12,14,6,13	5.1
	1,6,13,12,11	2,6,1,12,11	3.1
	1,6,11,12,13	*8,6,11,12,7*	*2.0*
	1,2,3,6,11	8,7,1,6,11	4.1
	6,7,11,12,13	13,12,2,1,11	6.1
	4,3,2,1,11	6,11,2,1,7	3.0
	1,6,7,8,9	**2,6,11,1,7**	**7.3**
	1,2,3,6,7	1,2,6,3,7	5.2

2	6,11,2,1,7	4,3,2,1,11	3.0
	1,11,12,13,14	3,2,1,7,6	6.1
	1,6,7,8,11	1,12,3,11,2	4.0
	2,1,6,12,11	2,7,6,1,11	5.2
	1,2,6,7,11	*1,2,3,12,11*	*2.0*
	8,7,6,1,11	11,12,13,1,6	5.1
	1,2,6,7,11	**1,6,12,11,2**	**6.2**
	1,6,7,11,12	**2,1,6,3,7**	**7.2**
	1,2,3,11,12	1,6,7,11,2	4.1
	11,1,6,12,13	11,1,6,7,12	3.1
	1,2,6,11,12	2,6,3,11,1	**7.3**
3	1,2,3,4,11	1,8,11,6,7	4.0
	1,11,12,13,14	*6,11,12,13,1*	*2.0*
	7,11,6,12,13	13,11,6,12,14	3.1
	1,2,3,6,7	**8,7,2,6,1**	**6.2**
	1,6,7,8,11	1,6,12,11,13	4.1
	8,1,7,6,9	11,8,7,6,9	5.2
	1,6,7,8,11	**13,12,11,7,6**	**7.2**
	1,11,6,12,2	12,11,6,7,8	3.0
	1,6,11,12,13	3,4,6,2,1	6.1
	1,2,3,4,5	1,9,8,6,7	5.1
	1,2,3,6,7	**1,6,7,2,3**	**7.3**
4	*6,7,11,12,13*	*6,2,11,12,1*	*2.0*
	6,7,8,9,11	**6,12,7,11,8**	**6.2**
	6,7,8,11,12	12,3,1,11,2	4.0
	2,1,7,6,11	4,1,2,3,5	5.1
	6,7,8,9,11	6,7,1,8,2	4.1
	3,11,6,1,2	3,11,2,1,12	5.2
	1,6,11,12,13	**7,12,11,8,6**	**7.3**
	1,2,3,4,11	1,2,12,3,11	3.1
	6,11,12,13,14	**11,12,8,7,6**	**7.2**
	1,2,6,11,12	8,12,11,7,6	6.1
	8,7,6,12,11	1,2,3,12,11	3.0